Proceedings of the 6th international conference on disability, virtual reality and associated technologies (ICDVRAT 2006)

The 6th International Conference on

Disability, Virtual Reality and Associated Technologies

Proceedings

Edited by:

Paul Sharkey (Programme Chair)
Tony Brooks (Conference Co-Chair)
Sue Cobb (Conference Co-Chair)

18 to 20 of September, 2006

Esbjerg, Denmark

ICDVRAT 2006

Proceedings:

The papers appearing in this book comprise the proceedings of the 6th International Conference on Disability, Virtual Reality and Associated Technologies, held on the 18th, 19th and 20th of September, 2006 in the Performance Arts Centre, Musikhuset Esbjerg, Denmark. The papers presented reflect the authors' opinions and are published as presented and without change (formatting and minor editing excepted). Their inclusion in this publication does not necessarily constitute endorsement by the editors, by ICDVRAT, or by the University of Reading.

Please use the following format to cite material from these Proceedings:

A.B. Author(s), "Title of paper", *Proc. 6th Intl Conf. on Disability, Virtual Reality and Assoc. Technologies*, in Sharkey, Brooks & Cobb (Eds.), page numbers, Esbjerg, Denmark, 18–20 Sept. 2006.

Proceedings reference number: ISBN 07 049 98 65 3

Published by

The University of Reading

For information, contact:
ICDVRAT, School of Systems Engineering (Cybernetics), University of Reading,
Whiteknights, Reading, RG6 6AY, UK

Phone: +44 (0) 118 378 6704 Fax: +44 (0) 118 378 8220

Email: *p.m.sharkey@reading.ac.uk* Web: *www.icdvrat.rdg.ac.uk*

Printed in the UK.

Contents

Abstracts & Information on Esbjerg

A Decade of Research in Disability and Virtual Reality

Session I Social Interaction

Session Chair: Tony Brooks

Session II Motion Tracking and 3D Modelling

Session Chair: Cecília Sik Lányi

Session III Therapy

Session Chair: Noomi Katz

Session IV Interaction Control

Session Chair: Bruno Herbelin

Session V Rehabilitation and Route Learning

Session Chair: Charles van der Mast

Session VI Medical Treatment and Home Based Rehabilitation

Session Chair: Pat Langdon

Session VII Visual Impairment

Session Chair: Ali Al-khalifah

 ISBN 07 049 98 65 3

Session VIII Cognitive Skills

Session Chair: Eva Petersson

Session IX Stroke Rehabilitation

Session Chair: Sue Cobb

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

Conference Organisation

Conference Co-Chairs

Tony Brooks, Aalborg University Esbjerg, Denmark
Sue Cobb, University of Nottingham, UK

Programme Chair

Paul Sharkey, University of Reading, UK

International Programme Committee

Jurgen Broeren, Göteborg University, Sweden
Jane Broida, Metropolitan State College of Denver, USA
Tony Brooks, Aalborg Universitet Esbjerg, Denmark
David Brown, Nottingham Trent University, UK
Sue Cobb, University of Nottingham, UK
Roy Davies, Lund University, Sweden
Judith Deutsch, University of Medicine & Dentistry of New Jersey, USA
Colin Harrison, University of Strathclyde, UK
William Harwin, University of Reading, UK
Maureen Holden, Northeastern University, Boston, USA
Noomi Katz, Hebrew University of Jerusalem, Israel
Tomohiro Kuroda, Kyoto University Hospital, Japan
Peter Kyberd, University of New Brunswick, Canada
Pat Langdon, University of Cambridge, UK
Rachel McCrindle, University of Reading, UK
Suzanne McDonough, University of Ulster, UK
Robin Morris, Institute of Psychiatry, London, UK
Eva Petersson, Aalborg Universitet Esbjerg, Denmark
John Rae, Roehampton University, UK
Albert (Skip) Rizzo, University of Southern California, Los Angeles, USA
Martin Rydmark, Göteborg University, Sweden
Jaime Sánchez, Universidad de Chile, Chile
Cecília Sik Lányi, University of Veszprém, Hungary
Sarah Sharples, University of Nottingham, UK
Penny Standen, University of Nottingham, UK
Penny Stribling, Roehampton University, UK
Danaë Stanton Fraser, University of Bath, UK
Daniel Thalmann, L'Ecole Polytechnique Fédérale de Lausanne, Switzerland
Cheryl Trepagnier, The Catholic University of America, Washington, DC, USA
Charles van der Mast, Delft University of Technology, Netherlands
Patrice (Tamar) Weiss, University of Haifa, Israel
Brenda Wiederhold, Virtual Reality Medical Center, USA
Paul Wilson, University of Hull, UK

Local Organising Committee

Tony Brooks, Aalborg Universitet Esbjerg, Denmark (Chair)
Jan Ciger, Aalborg Universitet Esbjerg, Denmark
Eric Dam, Erhvervsudvikling Vest, Esbjerg, Denmark
Bruno Herbelin, Aalborg Universitet Esbjerg, Denmark
Eva Petersson, Aalborg Universitet Esbjerg, Denmark
Jan Mols Poulsen, Tobakken, Esbjerg, Denmark
Paulina Ramirez, Aalborg Universitet Esbjerg, Denmark
Richard Sherwood, ICDVRAT Web Manager, University of Reading, UK
Christian Søvad, Musikhuset Esbjerg, Denmark
Susanne Theils, Aalborg Universitet Esbjerg, Denmark

Conference Series Archive

Paul Sharkey, University of Reading, UK
Richard Sherwood, ICDVRAT Web Manager, University of Reading, UK

Introduction

The purpose of the 6th International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT 2006) is to provide a forum for international experts, researchers and user groups to present and review how advances in the general area of Virtual Reality can be used to assist people with Disability.

After a peer review process, the International Programme Committee selected 34 papers for presentation at the conference, collected into 10 plenary sessions: Social Interaction; Motion Tracking and 3D Modelling; Therapy; Interaction Control; Rehabilitation and Route Learning; Medical Treatment and Home Based Rehabilitation; Visual Impairment; Cognitive Skills; and Stroke Rehabilitation. There is an additional session specifically for informal demonstrations, poster presentations and exhibits from a small number of companies.

The conference will be held over three days at the Musikhuset Performance Arts Centre, Esbjerg, Denmark.

ICDVRAT 2006 is the 10th anniversary of the first conference in the series, held in Maidenhead, UK (1996), with conferences held biennially in Skövde, Sweden (1998), Alghero, Sardinia, Italy (2000), Veszprém, Hungary (2002) and Oxford, UK (2004). To mark the 10th anniversary, the Proceedings begins with a review of the first decade of the conference. Abstracts from this conference and full papers from the previous conferences are available online from the conference web site *www.icdvrat.reading.ac.uk*. We are also pleased to be able to provide the complete ICDVRAT archive on CD-ROM with this Proceedings.

Acknowledgements

The Conference Chairs would like to thank the Programme Committee, for their input to the conference format and focus, and for their commitment to the review process, the authors of all the papers submitted to the conference, the Organisation Committee, and the students who help out over the period of the conference. On behalf of ICDVRAT 2006, we welcome all delegates to the Conference and sincerely hope that delegates find the conference to be of great interest.

Paul Sharkey, Tony Brooks and Sue Cobb

Conference Sponsors

ICDVRAT 2006 is sponsored by: the University of Reading, UK; SoundScapes, Denmark; and Aalborg University Esbjerg/Esbjerg Technical Institute, Denmark.

The conference is grateful for support from: Dronning Louise, Erhvervsudvikling Vest, Esbjerg Erhvervscenter, Esbjerg Turistkontor, Fiskeri- & Søfartsmuseet, Hotel Ansgar, Hotel Britannia, IBM, KPMG, Martin Professional, Musikhuset Esbjerg, Sony, Svømmestadion Danmark, Thuesen Bødker & Jæger, Tobakken, Vestjysk Marketing v/Bo Boysen, Videncenter Vest

Additional help in publicising the conference has been gratefully received from vrpsych-l@usc.edu, amongst many others.

Artwork

Artwork and Conference Layout by sckelp, adapted from original ideas of Eric Phipps and David Angus (ICDVRAT Logo) and from original clipart from the CorelDraw Clipart Library (ICDVRAT 2006 Graphic). Eric Phipps and David Angus (both of Project DISCOVER) and sckelp may be contacted through ICDVRAT.

Abstracts

<u>*In alphabetical order, based on first author*</u>.

Using virtual reality for medical diagnosis, training and education, **A Al-khalifah**, **R J McCrindle**, **P M Sharkey** and **V A Alexandrov**, University of Reading, UK

In this paper we present a number of the immersive VR applications that we have developed during the past 18 months as a means of practically demonstrating the modelling approaches previously reported. The paper discusses the usefulness of the different approaches in assisting medical practitioners to diagnose and track conditions which might lead to impairment or disability, and how they can be used to train medical students to recognise such conditions or to undertake associated medical procedures. Initial findings of a survey of undertaken with medical practitioners as to the effectiveness of VR and in particular immersive models as diagnostic and training aids are also presented.

Challenges in designing virtual environments training social skills for children with autism, **U Andersson**, **P Josefsson** and **L Pareto**, University West, SWEDEN

The purpose of the study is to explore particular challenges faced when designing virtual environments for children with autism, with the purpose of training social skills. Our findings are based on studying autistic behaviour during three years (primary and secondary sources), analysis of related system and other computer mediated assistive technology, as well as general game design. From these studies we have identified eight critical design parameters that need to be adjustable in a system suitable for autistic persons. The parameters importance, their variation range, as well as the need for *independent* adjustment of these were estimated and verified by experienced expert pedagogues.

Tongue-computer interface for disabled people, **L N S Andreasen Struijk**, Aalborg University, DENMARK

This work describes a new inductive tongue-computer interface to be used by disabled people for environmental control. The new method demands little effort from the user, provides a basis for an invisible man machine interface, and has potential to allow a large number of commands to be facilitated. The inductive tongue–computer interface implemented with 9 sensors was tested in three healthy subjects and the results shows typing rates up to 30 to 57 characters pr. minute after 3 hours of training.

Visual spatial search task (VISSTA): a computerized assessment and training program, **A Bar-Haim Erez**, **R Kizony**, **M Shahar** and **N Katz**, Hebrew University & Hadassah, Jerusalem/University of Haifa, ISRAEL

The aim of this paper is twofold 1) to introduce a computerized platform of Visual Spatial Search Task (VISSTA), its current package and potential for a variety of additional programs, and 2) to present results of the basic package of stroke patients and healthy controls. *Method*. Participants included 39 healthy individuals; 25 patients post right hemisphere damage (RHD) with unilateral spatial neglect (USN); 27 patients post RHD without USN; and 20 patients post left hemisphere damage (LHD). All participants were tested on the computerized VISSTA and paper and pencil cancellation tests. The stroke patients were also tested on the ADL checklist and FIM. *Results*. Findings indicate that the VISSTA is a valid visual search assessment that significantly differentiated between patients following stroke and healthy controls and between different stroke patient groups. USN patients showed impairment in both visual search conditions and clear laterality bias when target was presented on the left side of a computer screen, this was true for success rate and reaction time. RHD patients without USN performed better than those with USN, however, they still show impairment in attention properties of visual search and detection of targets (both on left and right) compared to healthy individuals. *Conclusions*. The VISSTA tool was found to be sensitive to levels of visual spatial attention by means of accuracy and reaction time. Results suggest that it is important to supplement the conventional paper and pencil tests and behavioral measures with tools that provide both accuracy and RT parameters in a randomized and more complex fashion. The VISSTA is also suitable for treatment as it provides a flexible platform.

A decade of research and development in disability and virtual reality and associated technologies: promise or practice, **S V G Cobb** and **P M Sharkey**, University of Nottingham/University of Reading, UK

The International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT) this year holds its sixth biennial conference and celebrates ten years of research and development in this field. A total of 180 papers have been presented at the first five conferences, addressing potential, development, exploration and examination of how these technologies can be applied in disabilities research and practice. The research community is broad and multi-disciplined, comprising a variety of scientific and medical researchers, rehabilitation therapists, educators and practitioners. Likewise, technologies, their applications and target user populations are also broad, ranging from sensors positioned on real world objects to fully immersive interactive simulated environments. A common factor is the desire to identify what the technologies have to offer and how they can provide added value to existing methods of assessment, rehabilitation and support for individuals with disabilities. We review this first decade of research and development in the ICDVRAT community and ask how far we have progressed: are we still discussing potential and promise or has our technology found its way into practical implementation?

Do we need high-scale flexibility in virtual therapies? **Z Geiszt**, **M E Kamson**, **C Sik Lányi** and **J A Stark**, University of Pannonia, HUNGARY/Austrian Academy of Sciences, AUSTRIA

Virtual reality (VR) offers a wide range of applications in the field of cognitive neuropsychology both in diagnosing cognitive deficits and in treating them. An optimal diagnostic method is the on-field test, which provides an opportunity to apply VR-based simulations. VR is also a useful tool for skill-building and training by setting up a virtual setting, which imitates the real environment including the attributes to be trained. Moreover, it provides a graded approach to problem-solving and the feeling of safety and it excludes the negative elements which are detrimental to the learning process. To further extend the effectiveness of VR applications it is necessary to refine VR environments and adapt them according to the specific needs of selected target groups and the real-time control of virtual events. Following the principle of flexibility we prepared two virtual environments: 1) an Adjustable Virtual Classroom (AVC) for the treatment of fear of public speaking in a primary school task-solving setting, and 2) a Virtual Therapy Room (VTR) designed for use with aphasic clients. Due to their flexible nature, a large number of elements can be customised in both of these settings including spatial organisation, textures, audio materials and also the tasks to be solved. The real-time control over the virtual avatars by the supervisor, i.e. therapist to guide the social interactions in the virtual world also allows him/her to follow-up on the user's reactions and therapy performance. By focusing on the details of the therapy room, we would like to demonstrate the relevance of the flexibility of the software in the development of innovative therapy solutions for aphasic clients.

Use of virtual reality as therapeutic tool for behavioural exposure in the ambit of social anxiety disorder treatment, **H Grillon**, **F Riquier**, **B Herbelin** and **D Thalmann**, EPFL, Lausanne, SWITZERLAND/ Aalborg University Esbjerg, DENMARK

We hereby present a study whose aim is to evaluate the efficiency and flexibility of virtual reality as a therapeutic tool in the confines of a social phobia behavioural therapeutic program. Our research protocol, accepted by the ethical commission of the cantonal hospices' psychiatry service, is identical in content and structure for each patient. This study's second goal is to use the confines of virtual exposure to objectively evaluate a specific parameter present in social phobia, namely eye contact avoidance, by using an eye-tracking system. Analysis of our results shows that there is a tendency to improvement in both the questionnaires and eye contact avoidance.

Tangible user interfaces: tools to examine, assess and treat dynamic constructional processes in children with developmental coordination disorders, **S Jacoby**, **N Josman**, **D Jacoby**, **M Koike**, **Y Itoh**, **N Kawai**, **Y Kitamura**, **E Sharlin** and **P L Weiss**, University of Haifa, ISRAEL/Osaka University, JAPAN/ University of Calgary, CANADA

Tangible User Interfaces (TUIs) are a subset of human-computer interfaces that try to capture more of the users' innate ability of handling physical objects in the real world. The TUI known as ActiveCube is a set of graspable plastic cubes which allow the user to physically attach or detach cubes by connecting or disconnecting their faces. Each cube is essentially a small computer which powers up and communicates with its neighbours upon connection to a neighbouring cube. When users assemble a physical shape using the system they also connect a network topology which allows ActiveCube to digitize and track the exact 3D geometry of the physical structure formed. From the user's perspective, ActiveCube is a very powerful tool; the 3D shape being built with it physically is tracked in the virtual domain in real-time. ActiveCube's use as a concrete, ecologically valid tool to understand dynamic functional processes underlying constructional ability in either typically developed children or in children with neurological pathology has not yet been explored. The objective of this paper is to describe the ActiveCube interface designed for assessing and treating children with Developmental Coordination Disorder (DCD). In our pilot study, six male children, aged 6 to 7 years, three with DCD and three who are typically developed were tested. The children's task was to successively use the ActiveCubes to construct 3D structures in a "matching" strategy known as "Perspective Matching". The usability results showed that all the participating children enjoyed the tasks, were motivated and maintained a high level of alertness while using the ActiveCubes. More than 80% of them found the tasks to be easy or moderate. "Similarity" data from single subjects has been used to show differences in constructional ability between children with DCD and those who are typically developed. This automated ActiveCube three-dimensional (3D) constructional paradigm has promise for the assessment and treatment of children with DCD.

Exploratory strategies and procedures to obtain non-visual overviews using TableVis, **J Kildal** and **S A Brewster**, University of Glasgow, UK

TableVis was developed to support computer users who are blind or visually impaired in tasks that involve obtaining quick overviews of tabular data sets. Previous work has covered the evaluation of this interface and its associated techniques of interactive data sonification and support exploratory processes. This paper examines the exploratory strategies and procedures employed by the users. A three-stage process for completing the exploratory task is described, and a discussion about the strategies and procedures that were observed is offered. Possible best practices and the most common issues are identified, which form the basis for the next steps to be taken in this line of research.

TheraGame – a home based VR rehabilitation system, **R Kizony**, **P L Weiss**, **M Shahar** and **D Rand**, University of Haifa, ISRAEL

The limitations of existing virtual reality (VR) systems in terms of their use for home-based VR therapy led us to develop 'TheraGame', a novel video capture VR system. TheraGame operates on a standard PC with a simple webcam. The software is programmed using a Java-based visual interaction system. This system enables a quick and easy definition of virtual objects and their behavior. The user sits in front of the monitor, sees himself and uses his movements to interact with the virtual objects. The objective of this presentation is to present the system, a number of the current applications, and some initial pilot usage results. Results from a study of 12 healthy elderly subjects showed moderate to high levels of enjoyment and usability. These scores were also high as reported by 4 participants with neurological deficits. Some limitations in system functionality were reported by one person with stroke who used TheraGame at home for a period of 2.5 weeks. Overall, TheraGame appears to have considerable potential for home based rehabilitation.

Constructing new coordinate system suitable for sign animation, **T Kuroda**, **K Okamoto**, **T Takemura**, **K Nagase** and **H Yoshihara**, Kyoto University Hospital/ Kyoto University, JAPAN

This paper proposes new coordinate system suitable for denoting sign language motion. As the proposed coordinate system consists of polar coordinate systems whose origins are certain points of human body, postures shown on the system can be proportional for avatars with any possible shape and fit with existing subjective sign notation systems. This paper extracted coordinate origins from Japanese-Japanese Sign Language Dictionary via morphological analysis. Selected 85 points are successfully mapped on H-ANIM standard humanoid avatar.

Hands-free man-machine interface device using tooth-touch sound for disabled persons, **K Kuzume** and **T Morimoto**, Yuge National College of Technology, JAPAN

This paper presents the realization of a hands-free man-machine interface using tooth-touch sound. The proposed device has several advantages, including low price, ease of handling, and reliability. It may be used as an Environmental Control System (ECS) and communication aid for disabled persons. We analyzed the characteristics of the tooth-touch sound, obtained using a bone conduction microphone. We then designed the device using VHDL (Hardware Description Language) and a simulation of the FPGA (Field Programmable Device) in practice. We applied our device to the ECS to demonstrate its usefulness and evaluate its performance. The results confirmed that the proposed device had superior features to comparable devices, such as those utilizing voice control or eye blinks, chin operated control sticks, mouth sticks, or a brain computer interface (BCI) for severely disable persons.

Combining interactive multimedia and virtual reality to rehabilitate agency in schizophrenia, **E A Lallart**, **S C Machefaux** and **R Jouvent**, Hôpital de la Salpêtrière, Paris, FRANCE

New interactive technologies offer the opportunity to involve the user's body in a virtual environment while seeing herself/himself performing the actions. Interactive exercises with a video-capture reinforce the perception-action loop, which is the pillar of agency (i.e. the ability to attribute the intention of an action to its proper author). We present a new paradigm as a possible treatment of agency disturbances in schizophrenia.

Statistical estimation of user's intentions from motion impaired cursor use data, **P Langdon**, **S Godsill** and **P J Clarkson**, University of Cambridge, UK

We report the application of new statistical state space filtering techniques to cursor movement data collected from motion impaired computer users performing a standard Fitts's Law style selection task. Developed as an alternative to expensive haptic feedback assistance, the aim was to assess the feasibility of the basic techniques in resolving the users intended trajectory from the extremely variable and wavering data that result from the effects of muscular spasm, weakness and tremor. The results, using a choice of basic parameter for the filters, show that the state space filtering techniques are well suited to estimating the intended trajectory of the cursor even under conditions of extreme deviation from the direct track and that these filters effectively act as an extreme cursor smoothing system. We conclude that further development of the approach may lead to more effective adaptive systems capable of providing smoothed feedback to the user and estimates of intended destination. A similar approach might further be applied to situationally induced movement perturbations.

Use of a virtual-reality town for examining route-memory, and techniques for its rehabilitation in people with acquired brain injury, **J Lloyd**, **T E Powell**, **J Smith** and **N V Persaud**, Birmingham University, UK

Route learning difficulties are a common consequence of acquired brain injury, and virtual environments provide a novel tool for researching this area. A pilot study demonstrated the ecological validity of a non-immersive virtual town, showing performance therein to correlate well with real-world route learning performance. The first patient study found that a rehabilitation strategy known as 'errorless learning' is more effective than traditional 'trial-and-error' methods for route learning tasks. The second patient study, currently in progress, will assess whether naturalistic route learning strategies of map and landmark use can be combined effectively with errorless techniques. A final study will investigate the relationships between route learning performance and scores on a select battery of neuropsychometric tests.

Exploring interpersonal dynamics between adults and motor disabled children within aesthetic resonant environments, **P Lopes-dos-Santos**, **L M Teixeira**, **S Silva**, **M Azeredo** and **M Barbosa**, University of Porto/Portuguese Catholic University, PORTUGAL

This paper focuses on interpersonal dynamics between the child with disabilities and the adult monitoring his/her performance in Aesthetic Resonance Environments. Drawing upon a social constructivist approach, a framework for human interactivity was checked against empirical data obtained from the exploratory implementation of an environment intending to stimulate body awareness and enhance movement in a group of six children with severe neuromotor disabilities. Results showed that the adult assumed the role of a facilitator, mediating interactions between children and the technological system. The provided social mediation increased quality of movement and improved levels of engagement in the observed group of participants.

The design of a haptic exercise for post-stroke arm rehabilitation, **E Lövquist** and **U Dreifaldt**, University of Limerick, IRELAND

In this paper we present an application based on an immersive workbench and a haptic device, designed to motivate stroke patients in their rehabilitation of their arm. The work presented in this paper is the result of a six-month project, based on evaluation with stroke patients and on informal interviews with medical doctors, physiotherapists and occupational therapists. Our application called "The Labyrinth" has been used for studies with stroke patients and we have seen that arm rehabilitation using Virtual Environments and haptics can be very encouraging and motivating. These factors are crucial to improve the rehabilitation process.

Brain-computer music interface for generative music, **E R Miranda**, University of Plymouth, UK

This paper introduces a brain-computer interface (BCI) system that uses electroencephalogram (EEG) information to steer generative rules in order to compose and perform music. It starts by noting the various attempts at the design of BCI systems, including systems for music. Then it presents a short technical introduction to EEG sensing and analysis. Next, it introduces the generative music component of the system, which employs an Artificial Intelligence technique for the computer-replication of musical styles. The system constantly monitors the EEG of the subject and activates generative rules associated with the activity of different frequency bands of the spectrum of the EEG signal. The system also measures the complexity of the EEG signal in order to modulate the tempo (beat) and dynamics (loudness) of the performance.

Tactile information transmission by apparent movement phenomenon using shape-memory alloy device, **Y Mizukami** and **H Sawada**, Kagawa University, JAPAN

This paper introduces the development of a tactile device using a shape-memory alloy, and describes the information transmission by the higher psychological perception such as the phantom sensation and the apparent movement of the tactility. The authors paid attention to the characteristic of a shape-memory alloy formed into a thread, which changes its length according to its body temperature, and developed a vibration-generating actuator electrically driven by periodic signals generated by current control circuits, for the tactile information transmission. The size of the actuator is quite compact and the energy consumption is only 20mW. By coupling the actuators as a pair, an information transmission system was constructed for presenting the apparent movement of the tactility, to transmit quite novel sensation to a user. Based on the preliminary experiment, the parameters for the tactile information transmission were examined. Then the information transmission by the device was tested by 10 subjects, and evaluated by questionnaires. The apparent movement was especially well perceived by users as a sensation of a small object running on the skin surface or as being tapped by something, according to the well-determined signals given to the actuators. Several users reported that they perceived a novel rubbing sensation given by the AM, and we further experimented the presentation of the sensation in detail to be used as a sensory-aids tactile display for the handicapped and elderly people.

Exploration of social rule violation in patients with focal prefrontal neurosurgical lesions, **R G Morris**, **E Pullen**, **S Kerr**, **P R Bullock** and **R P Selway**, Institute of Psychiatry, London/University of Nottingham/King's College Hospital, London, UK

Social rule violation was explored in 22 patients with prefrontal neurosurgical lesions and 22 normal controls. The patients were split into those with neurosurgical lesions impinging on the either the orbitofrontal (OF), dorsolateral (DL) or mesial (M) region of the prefrontal cortex. The study used a virtual reality 'bar' in which participants walked from the entrance to the bar counter, ordered drinks and returned to the entrance, with the choice of moving between other people (socially inappropriate) or around the people (social appropriate). There was a significant increase in socially inappropriate behaviour in the patients whose lesions were in other prefrontal regions than the dorsolateral prefrontal cortex.

Investigating the efficacy of a virtual mirror box in treating phantom limb pain in a sample of chronic sufferers, **C D Murray**, **E Patchick**, **S Pettifer**, **T Howard** and **F Caillette**, University of Manchester, UK

This paper describes a pilot clinical study to evaluate the efficacy of using immersive virtual reality (IVR) as a rehabilitative technology for phantom limb pain experienced by amputees. This work builds upon prior research which has used simple devices such as the mirror box (where the amputee sees a mirror image of their remaining anatomical limb in the phenomenal space of their amputated limb) to induce vivid sensations of movement originating from the muscles and joints of their phantom limb and to relieve pain. The IVR system transposes movements of amputees' anatomical limbs into movements of a virtual limb which is presented in the phenomenal space of their phantom limb. The primary focus here is on a qualitative analysis of interview data with each participant throughout the study. We argue that the findings of this work make a case for proof of principle for this approach for phantom pain treatment.

Perceptive 3-D interface via stereo observation, **D Padbury**, **R J McCrindle** and **H Wei**, University of Reading, UK

This paper describes an intuitive approach for interacting with a computer or computer-driven applications. Interaction is achieved by observing, through a stereo camera set-up, the motion of a user's hands. This motion is then translated into 3-dimensional (3-D) coordinates to enable interaction with either a traditional 2-dimensional (2-D) desktop or a novel 3-D user interface. The aim of this work is to provide an intuitive method of interaction to computer based applications for individuals whose condition might restrict their ability to use a standard keyboard/mouse.

Understanding users with reading disabilities or reduced vision: towards a universal design of an auditory, location-aware museum guide, **L Pareto** and **U Lundh Snis**, University West, Uddevalla, SWEDEN

We present ongoing work on the design of an information system for users with reading disabilities and users with reduced vision. The design target is a portable, auditory, location-aware information system, to complement visually displayed information in exhibitions. Applying a user-centred, we identify non-typical user-groups' specific requirements, which are turned into a design. The first design-iteration, which includes a formative evaluation, using a mock-up prototype, with dyslectic and visually impaired participants, is completed. The evaluation indicates that the user-group's specific aspects we have identified are relevant, while designing for these groups.

Evaluation of a computer aided 3D lip sync instructional model using virtual reality objects, **A Rathinavelu**, **H Thiagarajan** and **S R Savithri**, Dr Mahalingam College of Engineering and Technology/National Institute of Technology/All India Institute of Speech and Hearing, INDIA

Lip sync model is one of the aspects in computer facial animation. To create realistic lip sync model, a facial animation system needs extremely smooth lip motion with the deformation of the lips synchronized with the audio portion of speech. A deformable, parametric model of the lips was developed to achieve the desired effect. In order to create realistic speech animation, the articulatory modeling of the lip alone is insufficient. The other major articulators such as the tongue and jaw must also be considered. This lip sync model was initially developed by using polygonal model and blended key shape techniques and then parameterized by using 36 control points. The data for lip sync model was collected from video image and magnetic resonance imaging (MRI) techniques. The articulatory movements of our lip sync model were presented along with virtual reality (VR) objects in an interactive multimedia (IMM) interface. This IMM interface was used to teach small vocabulary of hearing impaired (HI) children. Virtual reality objects used to increase the cognitive process within HI children. The bilabial speech sounds were differentiated by using appropriate visual cues. Control panel was developed to present articulatory movements at different speed. For this study, six hearing impaired children were selected between the ages 4 and 7 and they were trained for 10 hours across 2 weeks on 18 meaningful words. The intelligibility of hearing impaired children was experimented to find out their performance in articulation and in memory retention. The results indicated that 65-75% of given words were articulated well and 75-85% of words were identified by all children.

User-centered design driven development of a virtual reality therapy application for Iraq war combat-related post traumatic stress disorder, **A A Rizzo**, **K Graap**, **J Pair**, **G Reger**, **A Treskunov** and **T Parsons**, University of Southern California/Virtually Better, Inc., Decatur, Georgia/U.S. Army, Fort Lewis, Tacoma Washington, USA

Post Traumatic Stress Disorder (PTSD) is reported to be caused by traumatic events that are outside the range of usual human experience including (but not limited to) military combat, violent personal assault, being kidnapped or taken hostage and terrorist attacks. Initial data suggests that at least 1 out of 6 Iraq War veterans are exhibiting symptoms of depression, anxiety and PTSD. Virtual Reality (VR) delivered exposure therapy for PTSD has been used with reports of positive outcomes. The aim of the current paper is to present the rationale, technical specifications, application features and user-centered design process for the development of a Virtual Iraq PTSD VR therapy application. The VR treatment environment is being created via the recycling of virtual graphic assets that were initially built for the U.S. Army-funded combat tactical simulation scenario and commercially successful X-Box game, *Full Spectrum Warrior,* in addition to other available and newly created assets. Thus far we have created a series of customizable virtual scenarios designed to represent relevant contexts for exposure therapy to be conducted in VR, including a city and desert road convoy environment. User-centered design feedback needed to iteratively evolve the system was gathered from returning Iraq War veterans in the USA and from a system in Iraq tested by an Army Combat Stress Control Team. Clinical trials are currently underway at Camp Pendleton and at the San Diego Naval Medical Center. Other sites are preparing to use the application for a variety of PTSD and VR research purposes.

Assisting the mobilization through subway networks by users with visual disabilities, **J H Sánchez** and **M A Sáenz**, University of Chile, Santiago, CHILE

We introduce AudioMetro, application software for blind users that represents a subway system in a desktop computer to assist mobilization and orientation in a subway network. A user can organize and prepare a travel by using the software before riding the subway. Conclusions of the usability study revealed the critical importance of using key interface elements, such as audio-based hierarchy menu, travel simulation, and information about the subway network, stations and their surroundings. Cognitive study results show an advance in the development of mobility skills needed for using the subway system which represent a contribution for a much more integral development of blind users and one-step towards social integration and inclusion.

Investigating the impact of method of immersion on the naturalness of balance and reach activities, **I Sander**, **D J Roberts**, **C Smith**, **O Otto** and **R Wolff**, University of Salford, UK

Immersive virtual reality is gaining acceptance as a tool for rehabilitation intervention as it places a person within a safe and easily configurable synthetic environment, allowing them to explore and interact within it through natural movement. The purpose of the study was to explore the usefulness of different types of virtual environments in the rehabilitation of upper limb function and balance in stroke patients. Although the above characteristics are ideal for rehabilitation of motor disorders, acceptance is hampered by insufficient knowledge of the effect of method of immersion on the naturalness of human movement. This study begins to address this by comparing the impact of two typical methods, Head Mounted Display (HMD) and immersive projection technology (IPT), on the naturalness of reach and balance activities. The former places the simulated image in front of the eyes, whereas the latter projects it around the user so that they perceive a holographic effect when wearing stereo glasses. Using the novel approach of placing the HMD in the IPT allowed subjects perceiving the environment through either, to be observed moving within the IPT holograph. Combined with sharing the same tracking and camera systems, this provided a direct comparison of tracking measurements, interaction behaviour, video and other observational data. The experiment studied subjects moving objects around a simulated living room setting initially on a level surface and then whilst varying the height and shape of the walking surface through raised planks. Performance in the synthetic environment, using both display types, was compared to that in a physical mock up of the living room. The experimental results demonstrate similar balance and reach movements in the physical mock up and the IPT, whereas a striking reduction in naturalness in both activities was observed for HMD users. This suggests that an inappropriate choice of method has the potential to teach unnatural motor skills if used in rehabilitation. Reasons for the difference are discussed along with possible remedies and considerations for practical applications within a clinical setting.

Development of vision based meeting support system for hearing impaired, **R Shikata**, **T Kuroda**, **Y Tabata**, **Y Manabe** and **K Chihara**, Kyoto University Hospital/Kyoto College of Medical Technology/ Nara Institute of Science and Technology, JAPAN

This paper describes a new meeting support system that helps the hearing impaired to understand the contents of the meeting. The proposed system distinguishes the mainstream of the discussion from other chattering based on the utterances. The situation of the meeting is acquired as a picture using an omni-directional vision sensor, and the system analyzes speaker's relations from the captured image by using face directions for the participants. The system shows the mainstream and the chattering of a meeting by using the analyzed result and speech-recognition.

Analysing the navigation of mentally impaired children in virtual environments, **C Sik Lányi**, **R Mátrai** and **I Tarjányi**, University of Pannonia, HUNGARY

In today's information society, computer users frequently need to seek for information on home pages as well as to select among software functions. A well-designed interface is essential in order to find everything necessary and meet the requirements of both the average user and users with special needs. Our project set out to discover where and with how much contrast objects should be placed on the screen in order to find everything easily. We examine what kind of characteristic searching routes can be found and whether we can find differences between the average user and mentally retarded user in navigation and everyday searching exercises.

Preliminary work for vocal and haptic navigation software for blind sailors, **M Simonnet**, **J-Y Guinard** and **J Tisseau**, European Center for Virtual Reality, École Nationale D'Ingénieurs de Brest, FRANCE

This study aims at the conception of haptic and vocal navigation software that permits blind sailors to create and simulate ship itineraries. This question implies a problematic about the haptic strategies used by blind people in order to build their space representation when using maps. According to current theories, people without vision are able to construct cognitive maps of their environment but the lack of sight tends to lead them to build egocentric and sequential mental pictures of space. Nevertheless, exocentric and unified representations are more efficient (Piaget et al, 1948). Can blind people be helped to construct more effective spatial pictures? Some previous works have shown that strategies are the most important factors in spatial performance in large-scale space (Tellevik, 1992) (Hill et al, 1993) (Thinus-Blanc et al, 1997). In order to encode space in an efficient way, we made our subject use the cardinal points reference in small-scale space. During our case study, a compass establishes a frame of external cues. In this respect, we support the assumption that training based on systematic exocentric reference helps blind subjects to build unified space. At the same time, this training has led the blind sailor to change his haptic strategies in order to explore tactile maps and perform better. This seems to modify his processing of space representation. Eventually, we would like to study the transfer between map representation and environment mobility. Our final point is about using strategy based on cardinal points and haptic virtual reality technologies in order to help the blind improve their spatial cognition.

Designing a device to navigate in virtual environments for use by people with intellectual disabilities, **P J Standen**, **D J Brown**, **N Anderton** and **S Battersby**, University of Nottingham/Nottingham Trent University, UK

As part of the process to design a device that would enable users with intellectual disabilities to navigate through virtual environments, an earlier study had collected baseline data against which to evaluate prototype design solutions. This study describes the evaluation of three design solutions: two modifications to a standard games joystick and a two handed device. Evaluation data were collected while 22 people with intellectual disabilities worked through four VE designed using a games format. None of the prototypes gave significantly improved performance over the standard joystick and some actually led to the user receiving more help from the tutor to use the device. This difference was significant for the two handed device in all four games. However there was considerable variation in results such that with some devices there was a reduction in the variability of scores between individuals. Future research needs to focus on the design of environments and how best to match the user with the device.

An evaluation of the use of a switch controlled computer game in improving the choice reaction time of adults with intellectual disabilities, **P J Standen**, **R Karsandas**, **N Anderton**, **S Battersby** and **D J Brown**, University of Nottingham/Nottingham Trent University, UK

The inability of people with intellectual disabilities to make choices may result from their lack of opportunities to practice this skill. Interactive software may provide these opportunities and software that requires a timed response may reduce choice reaction time. To test this, 16 people with severe intellectual disabilities were randomly allocated to either an intervention or a control group. The intervention group spent eight sessions playing a switch controlled computer game that required a timed response while the control group spent the same amount of time playing a computer based matching game that did not require a timed response. Both groups repeated a test of choice reaction time (CRT) that they had completed prior to the intervention. The intervention group made more accurate switch presses with repeated sessions while receiving less help from the tutor who sat alongside them. The intervention group also showed a significant reduction in their CRT from baseline while the control group did not.

Investigating the use of force feedback joysticks for low cost robot-mediated therapy, **H Sugarman**, **E Dayan**, **A Lauden**, **A Weisel-Eichler** and **J Tiran**, Ono Academic College/Hadassah College Jerusalem/ Ben-Gurion University of the Negev, ISRAEL

We are developing a low-cost robotic system– the Jerusalem Telerehabilitation System – using a force feedback joystick and a standard broadband internet connection. In this study, the system was found to be user-friendly by therapists, older and younger normal subjects, and a post-stroke subject. Kinematic analysis of the joystick movements showed differences between the older and younger normal subjects and between the post-stroke subject and older normal subjects. These preliminary data indicate that our low-cost and straightforward system has the potential to provide useful kinematic information to the therapist in the clinic, thereby improving patient care.

Developing an ENABLED adaptive architecture to enhance internet accessibility for visually impaired people, **C C Tan**, **W Yu** and **G McAllister**, Queen's University Belfast, UK

This paper gives an overview of the current status of Internet accessibility and offers a brief review of the existing technologies that address accessibility problems faced by visually impaired people. It then describes an adaptive architecture which is able to integrate diverse assistive technologies so as to allow visually impaired people to access various types of graphical web content. This system is also capable of adapting to user's profile and preferences in order to provide the most adequate interface to the user.

Virtual social environment for preschoolers with autism – preliminary data, **C Y Trepagnier**, **M M Sebrechts**, **A Finkelmeyer**, **J Woodford** and **W Stewart Jr**, The Catholic University of America, Washington, DC, USA

Preliminary results are presented of a feasibility study, still in progress, of a virtual social environment designed to stimulate the social attention of pre-school-aged children with Autism Spectrum Disorder (ASD). The system uses eye-tracking and provides gaze-contingent rewards of clips from preferred videos. Of six children reported on here, most find the experimental setting appealing, and the rewards compelling; they voluntarily engage with it across numerous sessions, and demonstrate learning, with large inter-individual differences in rate of progress. Implications are discussed for the pilot study to follow.

Technological challenges and the Delft virtual reality exposure system, **C A P G van der Mast**, Delft University of Technology, THE NETHERLANDS

In this paper the architecture and use of the Delft VRET system is described. In this generic VRET system special emphasis is given to usability engineering of the user interface for the therapist. Results of controlled experiments with patients are summarized. The system is in regular use in a few clinics since 2005. New technological and functional challenges of VRET are presented. These challenges will lead to improvements of the system in the future. Agent support for the therapist and tele-VRET are the most promising challenges.

Virtual reality for interactive binocular treatment of amblyopia, **P E Waddingham**, **S V G Cobb**, **R M Eastgate** and **R M Gregson**, University of Nottingham, UK

Amblyopia, or ‘lazy eye’, is currently treated by wearing an adhesive patch over the non-amblyopic eye for several hours per day, over a period of many months. Non-compliance with patch wearing is a significant problem. Our multi-displinary team involved clinicians and technologists to investigate the application of VR technology in a novel way. We devised a binocular treatment system in which children watch a video clip of a cartoon on a virtual TV screen, followed by playing an interactive computer game to improve their vision. So far the system has been used to treat 39 children of which 87% have shown some improvement in vision. Vision improvement tended to occur within the first 3-4 treatment sessions. This paper describes research development of the I-BiT™ system. We present a summary of results from clinical case studies conducted to date and discuss the implications of these findings with regard to future treatment of amblyopia.

Non-visual feedback for pen-based interaction with digital graphs, **S A Wall** and **S A Brewster**, University of Glasgow, UK

Access to simple visualisations such as bar charts, line graphs and pie charts is currently very limited for the visually impaired and blind community. Tangible representations such as heat-raised paper, and inserting pins in a cork-board are common methods of allowing visually impaired pupils to browse and construct visualisations at school, but these representations can become impractical for access to complex, dynamic data, and often require a sighted person’s assistance to format the representation, leading to a lack of privacy and independence. A system is described that employs tactile feedback using an actuated pin-array, which provides continuous tactile feedback to allow a visually impaired person to explore bar charts using a graphics tablet and stylus. A study was conducted to investigate the relative contribution of multimodal feedback (tactile, speech, non-speech audio) during typical graph browsing tasks. Qualitative feedback showed that the participants found it difficult to attend to multiple sources of information and often neglected the tactile feedback, while the speech feedback was the most popular, and could be employed as a continuous feedback mechanism to support graph browsing.

Evaluation of virtual reality therapy in augmenting the physical and cognitive rehabilitation of war veterans, **B K Wiederhold** and **M D Wiederhold**, Interactive Media Institute/Virtual Reality Medical Center, San Diego, CA, USA

War veterans with neuromuscularskeletal injury often require significant treatment and rehabilitation, straining health care resources. In a study funded by the Office of Naval Research (ONR), the Virtual Reality Medical Center (VRMC) is applying virtual reality therapy to injured military personnel at the Naval Medical Center San Diego (NMCSD). The goal of this study is to investigate whether augmenting traditional rehabilitation with VR (in this case, off-the-shelf interactive video games) will enable a more rapid and complete rehabilitation. Because VR is interactive, and encourages patients to use their entire body to reach goals in the game, it is conceivable that it will make rehabilitation sessions more comfortable and entertaining. Participants consist of 20 veterans with upper arm and shoulder injuries (rotator cuff tear, shoulder impingement, bursitis) or amputation. The participants are divided into two groups (n=10): an experimental group, which receives traditional rehabilitation augmented by virtual reality therapy, and a control group, which undergoes traditional rehabilitation. Participants will complete ten treatment sessions in their respective condition. Though the study has not been completed, preliminary results based on subjective questionnaires and functional capacity indicate that the experimental condition may elicit increased heart rate and respiration. Participants in this group also seem to enjoy the music and interaction made possible through VR. These results suggest that VR may enhance the rehabilitation process, creating a more effective form of treatment. Long-term benefits of this form of treatment may include improved treatment time and reduced drop out rates, therefore reducing the costs of rehabilitation.

SMART project: application of emerging information and communication technology to home-based rehabilitation for stroke patients, **H Zheng**, **R Davies**, **H Zhou**, **J Hammerton**, **S J Mawson**, **P M Ware**, **N D Black**, **C Eccleston**, **H Hu**, **T Stone**, **G A Mountain** and **N D Harris**, University of Ulster/ University of Essex/Sheffield Hallam University/University of Bath, UK

The SMART project, entitled 'SMART rehabilitation: technological applications for use in the home with stroke patients', is funded under the EQUAL (extend quality of life) initiative of the UK Engineering and Physical Sciences Research Council (EPSRC). The project aims to examine the scope, effectiveness and appropriateness of systems to support home-based rehabilitation for older people and their carers. In this paper, we describe the design and development of a low-cost home-based rehabilitation system. Through the project we have involved end users in the design process and this model can be applied to the design of other healthcare related systems.

Esbjerg – Gateway to Scandinavia

Tony Brooks

Aalborg University Esbjerg/Esbjerg Technical Institute, Denmark

Welcome to Esbjerg, Denmark's 5th largest but youngest city which, in August, was crowned as the Danish "City of the Year".

Esbjerg is located in Jutland, which is the Danish mainland that borders Northern Germany. Upon experiencing the region's unspoiled natural beauty that is exemplified through its clear flowing waterways, secluded bays and inlets, and the expansive white sand beaches that are fringed with emerald green forestry and rolling countryside, one can easily understand why Esbjerg is referred to as the jewel on the west coast and the 'Gateway to Scandinavia'.

In the mid nineteenth century when King Christian the 9th declared that a harbour be built in Esbjerg to satisfy the growing agricultural exports, little could he have believed the magnitude his foresight would one day contribute to the region. This realised when, in celebrating its centennial in 1968, the city luminaries acknowledged the harbour as the region's main industrial hub. Today, the harbour hosts the international DFDS passenger ferry terminal and related industries that are channelled from the North Sea; typically fish and energy, the latter being specific to the growing oil exploration rigs and offshore wind farms that are serviced from the city and its airport.

Departing the harbour northwards along the coast road one arrives at the Esbjerg memorial, a statue made from granite that acts as a harsh reminder of the strength and power of nature and the inherent dangers of the sea. The memorial commemorates the community's lives lost at sea since 1900. This memorial's neighbour, the nine-meter-tall sculpture "Man Meets the Sea", created by Svend Wiig Hansen in 1995, depicts man gazing out into the magnificence of nature.

Opposite the Wiig Hansen monument is the Fisheries and Maritime Museum. The museum was opened in 1968 in celebration of the harbour's centennial. This venue offers an illustrated history of fishing and shipping in the region, while in the large outdoor exhibition it is possible to experience a real harbour environment. The museum also contains a large saltwater aquarium, a 'sealarium' (the seals are fed at 11:00 and 14:30), and a live mink enclosure.

A few kilometres further north is Hjerting Church, which boasts an altar decoration from 1992 by Robert Jacobsen that tells the story of Christ's Passion with both simplicity and empathy. The architecture is also interesting, since it combines the medieval church tradition with a modern conception of space. Additional alternatives are the twelfth century Jerne Church and the thirteenth century Guldager Church, which offer original gothic interior decorations.

At the West Coast road is Eva Koch and Steen Høyer's interactive work titled 'Light Mound' (1997), a colossal mound of earth with tiny cupolas of light spread out over it. It is Denmark's largest work of art, 180 metres in diameter, with an adjacent earthwork of 320 metres that is intersected by a road. The 'Light Mound' will not be complete until it is clad with heather, a process that will take years. The pulsating of the light is best seen at night – the light intensity and rhythm being determined by the intensity of the traffic on the West Coast road.

The oldest town in Denmark is Ribe, which lies thirty kilometres south from Esbjerg. Famous as a central point from which to explore Viking folklore and history, Ribe, in addition to its Viking museum, hosts a summer Viking market and village scene (including tradesmen, artists, huntsmen and warriors' battles) at the authentic Viking Centre which lies on the outskirts of the town.

Ribe Art Museum was inaugurated in 1891 and has a valuable collection of work dating from 1790. The museum is housed in one of the most important private buildings in Ribe: the factory owner Balthazar Giørtz' large private establishment, built 1860-64 after drawings made by the royal surveyor L A Winstrup. The collections show the main line of Danish pictorial art from c. 1750 to 1940 with masterpieces by Jens Juel, C.W. Eckersberg, Christen Købke, Kristian Zahrtmann, L.A. Ring, P.S. Krøyer, Anna and Michael Ancher, William Scharff, amongst others.

Perhaps you get the feeling that you are surrounded by art? Indeed, the region proudly boasts a reputation as a cultural Mecca that is embedded with numerous art galleries, museums and private studios. Prominent amongst the traditional on offer in the city is the Esbjerg Museum where, alongside historic recreations of city shops and streets from the 19th century are Iron and Viking Age exhibits. The museum also contains Denmark's finest amber museum, which illustrates the history of amber over a period of 10,000 years.

For the more contemporary in taste the Esbjerg Art Museum, housed in the same building as the conference venue Musikhuset, should satisfy even the most discerning. It is a collection of approximately 800 works that include Danish art produced from around 1920 up to the present day, by names such as the COBRA painters Richard Mortensen, Robert Jacobsen, Svend Wiig Hansen and Per Kirkeby, and by representatives of the latest trends in art including Christian Lemmerz, Michael Kvium and Peter Bonde. The sculptured sign "Esbjerg" residing outside the museum is a spectacular example of Robert Jacobsen's work from 1963. The museum features two novelties in the museum world: open storehouses and an aesthetics laboratory. In the open storehouses paintings that are not on display in the permanent exhibition can be seen on forty pullout walls. This allows visitors to create their own exhibition selection. Visitors can also experience the aroma installation, ("Wittgenstein's Garden") by Skeel & Skriver, a work that challenges traditional notions of what art is all about. In the aesthetics laboratory, groups can experiment with various aesthetic phenomena.

Adjacent to the Esbjerg Art Museum is the nineteenth century Water Tower (1896−7) that was designed by the local architect C. H. Clausen, whose many other buildings may be seen in the city of Esbjerg. It is uniquely designed in the theme of a noble German residence of the Middle Ages including an imposing battlement. As well as its stunning exterior, the tower offers some of the best views of the harbour and Esbjerg as well as contributing its own art exhibitions.

Overseeing the bustling central square, where most visitors to Esbjerg congregate to shop along adjacent walking streets or to enjoy the relaxing café atmospheres, proudly stands the equestrian statue of Christian IX – king of Denmark when Esbjerg was founded in 1868. On the distinctive corner of the square lies the former courthouse and county jail, which now houses Esbjerg Tourist Office. Across from the Tourist Board is the Henry Heerup garden featuring a number of his granite works. Heerup was also a member of the COBRA group with featured exhibits at leading museums around the world.

A short walk further east is the West Jutland Academy of Music is in a converted electricity factory that now houses the city's latest work of art, Thorbjørn Lausten's "The Clock" (1998); this combines image and sound in a rigorous composition that is governed by the aesthetic principles of geometry.

In recent years the heart of the city has a clearly detectable energy directed towards an ambition to evolve, achieve and lead. Two universities are amongst eleven higher educational institutes supporting the city's ambition to further develop its profile as a nucleus for knowledge exchange, innovation and entrepreneurship. The city has established an innovations centre with professional advisors where the graduated students can go and get advice, and establish their own company after leaving the University, with financial support being provided in the start up phase.

The thriving business community is made up of companies with international customer databases. Such global networking reflects the city's youth and the ambition of a forward momentum towards an identity with inherent contemporary vision. The city's cultural infrastructure substantiates the vision with substantial investment that is exemplified by Esbjerg possessing one of the leading venues in the country for conferences, exhibitions and entertainment – namely the Utzon designed Esbjerg Performing Arts Centre, Musikhuset (1997). This vision is also reflected through the top international celebrities that are engaged to perform here: in 2006 the featured artist was Placido Domingo who performed to a sell out international audience in July.

Culture is never far away in Esbjerg as theatre, dance and live music can be found on most nights in the city venues, clubs, bars, and selection of fine cosmopolitan restaurants. Also, give a moment to meet the congenial natives of Esbjerg who have a reputation as being amongst the most "Hygge" in Denmark.

My first experience of Esbjerg was when my exhibitions toured leading Scandinavian Museums of Modern Art in the late 1990's. During my stay I remember being struck by the city's abundance of art and the community's interest in promoting their artisans. Esbjerg and the people left its mark on me and I am pleased to have returned to contribute in whichever way I can to assist in promoting the city, develop the profile and to educate the next generation of artist innovators.

I hope truly that your experience of Esbjerg approaches mine, and that you, too, return.

Acknowledgements. I am grateful for the invaluable assistance of Lise Retbøll, Tourist Board, Esbjerg, in compiling and checking the details in this essay.

ICDVRAT 2006

A Decade of Research in Disability and Virtual Reality

Chair: Tony Brooks

A decade of research and development in disability, virtual reality and associated technologies: promise or practice?

S V G Cobb[1], and P M Sharkey[2]

[1]VIRART, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

[2]University of Reading, Reading, UK

sue.cobb@nottingham.ac.uk, *p.m.sharkey@reading.ac.uk*

[1]*www.virart.nottingham.ac.uk*, [2]*www.cyber.rdg.ac.uk*

ABSTRACT

The International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT) this year holds its sixth biennial conference and celebrates ten years of research and development in this field. A total of 180 papers have been presented at the first five conferences, addressing potential, development, exploration and examination of how these technologies can be applied in disabilities research and practice. The research community is broad and multi-disciplined, comprising a variety of scientific and medical researchers, rehabilitation therapists, educators and practitioners. Likewise, technologies, their applications and target user populations are also broad, ranging from sensors positioned on real world objects to fully immersive interactive simulated environments. A common factor is the desire to identify what the technologies have to offer and how they can provide added value to existing methods of assessment, rehabilitation and support for individuals with disabilities. We review this first decade of research and development in the ICDVRAT community and ask how far we have progressed: are we still discussing potential and promise or has our technology found its way into practical implementation?

1. INTRODUCTION

The first European Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT) was held in the UK in 1996. 30 papers were presented. Keynote addresses described the increasing interest in virtual reality technology for disabilities (Murphy, 1996), potential application in neurological rehabilitation (Rose, 1996) and as a means of providing access to computer information for users with visual impairment (Zwern and Goodrich, 1996). The second ECDVRAT was held in Sweden in 1998, establishing the biennial timing of the conference, whilst the international attendance in '96 and '98 prompted a change in title from 'European' to 'International' for the 2000 conference, held in Alghero, Sardinia. A chance contribution of an essay on the coastal region of Alghero led to a now traditional essay on the host town for each conference since: Veszprém (Hungary, 2002), Oxford (UK, 2004) and Esbjerg (Denmark, 2006).

As an applied research area, the ICDVRAT community includes practitioners, educators, researchers, technologists and end users from schools, hospitals, disability service providers, rehabilitation institutes academic research, scientific institutes and technology development labs drawn from a variety of disciplines including; medicine, healthcare, education, computer science, psychology and engineering. Papers presented at ICDVRAT describe technology development, design, evaluation and impact of virtual reality and associated technologies via individual case studies, experimental studies and large scale multi-centre research projects.

This paper celebrates 10 years of ICDVRAT by presenting a representative review of papers presented over the past decade, organised around three central themes:

- Virtual Reality and Associated Technologies – what are they?
- Disability – for what user populations have these technologies been developed?
- Application usage – how has the technology been applied?

We conclude by examining the ICDVRAT literature to see what changes have occurred over the last decade and pose the question: is this a technology in practice or are we still just offering promise?

For reasons of brevity, key aspects of technology development and application are described only. We apologise to authors whose papers are not represented in this review, it was not possible to include them all.

2. VIRTUAL REALITY AND ASSOCIATED TECHNOLOGIES

2.1 3D Virtual Environments

The predominant technology applied in ICDVRAT is Virtual Reality (VR) or Virtual Environments (VEs). These are computer-generated three-dimensional environments that can be explored and interacted with in real-time. The most commonly–used VE development platform is Superscape™ and this has mostly been used to construct representations of pseudo and actual real-world environments for training, education and performance assessment purposes. Other VE development platforms used are: dVISE, VRML, WorldUp, World Toolkit, HalfLife, Macromedia Director and Shockwave.

These VEs can be displayed via standard desktop PC delivery, Head-Mounted Displays (HMD), single screen projection, or CAVE™–type multiple screen projection systems, an example of which is the University of Reading's ReaCTor, as illustrated on the back cover of this Proceedings. The majority of virtual environments have been presented via desktop displays although some researchers have used the specific advantages of HMD and projected displays to provide an 'immersion' experience. More bespoke display systems have also been presented such the Immersadesk and the University of Southern California's panoramic display suite. While use of such systems is limited by budget a recent theme has developed in the use of simple technology and widely available toy/game interfaces such as the EyeToy for rehabilitation programmes.

2.2 Multimedia

Virtual environments do not need to be 3D and 2D projected environments can also provide a sense of 'immersion' or 'engagement'. For some applications, such as spatial navigation training of real places, video images of the real world are more appropriate and a lot easier to generate than computer simulation models. The level of video capture has ranged from full 360° panorama scenarios with virtual characters inserted into the scenes for anger management and treatment of social phobias (e.g. Rizzo et al., 2004), the VividGroup's Gesture Xtreme (GX) system, to the use of the EyeToy and simple video capture to develop very low cost systems for physiotherapy following stroke (e.g. Rand, Kizony and Weiss, 2004).

2.3 Multi-sensory and acoustic environments

Virtual environments do not need to be visual. Considerable research has been conducted to develop acoustic environments. First presented as a concept design in 1996, it was not clear whether it was possible to render 3D sound in real time (Lumbreras, Barcia and Sánchez, 1996). The existing acoustic technology formats were not appropriate for 3D surround sound to enable accurate perception of sound direction and Keating (1996) demonstrated use of the ambisonic-B format. In 1998, Lumbreras and Sánchez presented results showing that spatialised sound could be used to stimulate diminished cognitive skills in blind children and can be further used to assist in navigation.

Winberg and Hellström (2000) demonstrated the possibility of auditory direct manipulation with the sonified Towers of Hanoi task. This was promising evidence that acoustic environments could provide access to computers for blind users, although this had not been tested with blind users. The spatial audio system (Kurniawan et al., 2004) demonstrated that blind users could differentiate sounds representing sound effects created in a small, medium or large real or virtual room, indicating that the sound algorithms were appropriate.

Tony Brooks et al. (2002) showed that the term 'virtual environments' need not be restricted to a limited notion of an architecturally understandable space, but can be realised as a more visually and sonically abstract space that can enhance quality of life for severely disabled children.

2.4 Interaction methods

Different technologies have been used to provide interaction with virtual environments. These include standard PC interaction devices (keyboard, mouse and joystick), VR specific interaction devices for interaction and navigation with a 3D virtual environment (data gloves, HMD and tracker systems), and specialised technologies developed or adapted for ICDVRAT users. Examples include augmented reality systems to integrate technologies for home-based rehabilitation (Hammond, Sharkey and Foster, 1996), eye-tracking technology for gaze-control interaction (Istance et al., 1996; Bates and Istance, 2004), whole body movement, and tangible interfaces. Tangible interfaces have included navigation controlled via an exercise bike (Johnson, Rushton and Shaw, 1996), a wheelchair (Harrison et al., 2000), and tethered kitchen items (Hilton et al., 2002).

Gesture recognition technologies have been developed using cyberglove and position sensors for sign language input (Vamplew, 1996; Kuroda, Sato and Chihara, 1998; Sawada, Notsu and Hashimoto, 1998) and camera-vision systems for either modelling of hand-gesture for communication (García-Ugalde, Gatica-Pérez and García-Garduño, 1998) or for interaction with the real world (Foyle and McCrindle, 2004; Pridmore et al., 2004).

2.5 Haptic, force-feedback and tactile devices

Tactile feedback to users has been explored using tactile tablets (Eves and Novak, 1996; Sánchez and Flores, 2004), the Impulse Engine force-feedback device (Colwell et al., 1998), a vibrotactile actuator (Langdon et al., 2000), force-feedback mouse (Langdon et al., 2002; Caffrey and McCrindle, 2004) and force-feedback joysticks (Conner et al., 2002; Sánchez and Flores, 2004).

The most explored haptic device has been the PHANToM haptic stylus used for tactile exploration of 3D data (Jansson, 1998; Jansson 2000; Petrie et al., 2000, Wall and Brewster, 2004) and assessment and rehabilitation of fine manual dexterity (Broeren et al., 2002). The InMotion2 Robot with graphic display was also developed for motor control, providing tele-assessment and co-operative rehabilitation (in which the patient's robot mimics the clinician's robot) of hand/arm movement (Olsson, Carignan and Tang, 2004).

2.6 Wheelchair-mounted devices

A range of sensors have been mounted onto wheelchairs to provide navigation feedback and obstacle detection. These include ultrasonic sensors (Gunderson, Smith and Abbott, 1996), vision, GPS and GIS (Mori and Kotani, 1998) and the MANUS manipulator (ten Kate et al., 2000).

3. DISABILITIES

3.1 Visual impairment

Given that 'virtual reality' has, since its earliest development, concentrated primarily on the presentation of a high fidelity visual experience, it is perhaps surprising that the largest single group of users mentioned in the ICDVRAT proceedings are those with visual impairment.

An early proposal for use of virtual reality and associated technologies for users with visual impairment was the screen enlarger software with HMD delivery (Zwern and Goodrich, 1996). A variety of applications for this technology for blind users was proposed (Cooper and Taylor, 1998) and many papers describe application of tactile feedback technology (described earlier in §2.5) to enable visually impaired users to explore 3D data. Early studies demonstrated the potential (Jansson, 1998) as well as the problems (Colwell et al., 1998) and improved design resulted in successful perception of textures using haptic interfaces (Petrie et al., 2000). Tactile information can be enhanced by auditory and visual information although it was found that pictorial information must be simplified for haptic reading (Jansson, 2000). Consideration must also be given to ensure that the data exploration software does not in itself place increase memory demands upon the user, such as the need to remember 'beacons' marking points of interest in numerical data (Wall and Brewster, 2004).

Acoustic virtual environments have been used for spatial mapping and navigation. After several years of early development (Lumbreras and Sánchez, 1998; Berka and Slavik, 1998; Lahav and Mioduser, 2000-2), later studies found that, after exploring in virtual environments, children then explored a real environment more quickly and confidently (Sanchez et al., 2000; Lahav and Mioduser, 2004; Feintuch et al., 2004), indicating that they had been able to construct a mental model of the environment. Other research found that multimodal interaction, combining haptic and auditory information, enhanced access to 3D computer images (Gladstone, Graupp and Avizzano, 2002) and could be used for internet navigation (Caffrey and McCrindle, 2004).

The impact of acoustic-virtual environment exploration on cognitive learning for children with visual impairment has also been examined. Sánchez and Lumbreras (2000) found that learning mental structure through sound is possible and that sighted children do not use the same learning method as visually impaired children. Further study demonstrated that blind children could construct mental images of 3D space, and showed evidence of improvement in haptic perception, abstract memory, and spatial abstraction (Sánchez, 2004). Involving blind children in design of an audio-based interactive interface for learning and cognition of maths concepts was successful; the AudioMath programme resulted in memory and knowledge improvement and the system was highly accepted by users (Sánchez and Flores, 2004).

3.2 Cognitive Impairment

Several research groups have examined potential use of virtual environments for assessment and rehabilitation of patients with acquired cognitive disability due to stroke or brain injury.

Pugnetti et al. (1996) examined nervous system correlates of the virtual reality experience and found that it was possible to monitor electrophysiological brain activity whilst using VR. Other studies have also found that physiological measures can be used to assess patient responses to virtual environments (Meehan et al., 2000; Herbelin et al., 2004). Experimental studies demonstrated effective assessment of aspects of spatial memory in patients with neurological impairment (Pugnetti et al., 1998, 2000) and studies with healthy undergraduate students demonstrated superior recall of spatial layout when subjects were engaged in active participation and superior object memory in passive participation (Attree et al., 1996). Comparison of performance in virtual environment versus real world prospective memory tasks between stroke patients and healthy matched adults identified memory impairment in the stroke patients (Brooks et al., 2002). This method was successfully applied to the assessment of executive functioning, including prospective memory, in patients with prefrontal lesions (Morris et al., 2002a). Studies using immersive virtual environments were also found to show clear spatial memory deficit in patients with right temporal lobectomy (Morris, Parslow and Reece, 2000) and allocentric spatial memory in patients with anoxic hippocampal damage (Morris et al., 2002b).

3.3 Motor impairment

Research for this group generally falls into two categories: development of interaction methods providing users with access to computers (detailed in §4.1), or use of virtual reality and associated technologies for assessment or rehabilitation of motor control.

An early concept demonstration of the use of blue screen technology to provide patients with a video image of themselves on a television screen interacting with virtual objects in a video game was presented by Joyce and Phalangas (1998). This concept was extended to large project screen for gross motor rehabilitation of patients with stroke, spinal cord injury and cerebral palsy. Pilot studies using the Vivid system were encouraging, with evidence of user involvement in the VE and increased mobility (Kizony et al., 2002) and a later experimental study with stroke patients established measures of presence and perceived exertion (Kizony, Katz and Weiss, 2004). Further study compared use of the VividGroup's Gesture Xtreme (GX) system with a Sony PlayStation II EyeToy (Rand, Kizony and Weiss, 2004). Elderly patients preferred the EyeToy system. However, although the EyeToy offers considerable cost advantage over the GX system, the software games are not written for this purpose and therefore offer limitations for rehabilitation.

A similar theme has been followed by Smyth and Wann (2000) in the use of low cost interactive interfaces for movement rehabilitation. Using off-the-shelf force feedback joysticks, such as Microsoft Sidewinder or the Logitech Wingman, to develop simple navigation tasks for reinforcement learning in patients suffering from stroke. Limitations of motion provided by joysticks motivated alternative solutions such as those provided in Louriero, Collin and Harwin (2004) who have augmented the interaction with the virtual task through the provision of a large reach robot to give robot assisted motion therapy.

VRAT has been applied to rehabilitation of fine motor control (Crosbie et al., 2004). Wearing a data glove, patients performed hand movement tasks involving wrist extension and reach and retrieve tasks with visual feedback of performance provided by the VE. Results were generally favourable and patients did report experiencing fatigue and exertion, reflecting the increase in motor activity demanded by the tasks.

3.4 Learning disability

Virtual environments providing task-based training and education in everyday living or vocational skills were proposed in 1996 (Brown and Stewart, 1996). Successful transfer of learning from use of a virtual supermarket to the real world was demonstrated (Cromby et al., 1996) and transfer of learning and increased engagement in the task was found following virtual environment training in travel, shopping and ordering food in a café (Cobb, Neale and Reynolds, 1988). Evidence of transfer of training was also found following virtual environment training in kitchen skills for students with learning disabilities attending a catering college (Rose et al., 1998; Rose, Brooks and Attree, 2000) and the VIRT factory trainer project developed a commercially available training package for users with learning difficulties seeking employment in sheltered factories (Mendozzi et al., 2000). Whilst successful transfer of training was demonstrated in these projects, it was suggested that virtual environment training is not *better than* other training methods for these users (Rose, Brooks and Attree, 2000). Furthermore, other studies have found that students require tutor instruction to guide them through virtual environment interaction (Standen and Low, 1996; Standen et al., 2000) and that the virtual reality training module should be incorporated as part of a larger training programme (Shopland et al., 2004).

Some studies have specifically examined user control over virtual environment interaction finding that, with tutors proving support (‘scaffolding’), students do progressively make more self-initiated actions and therefore gradually increase self-directed learning (Standen and Low, 1996). Studies examining use of interaction devices found that dual control devices are confusing for users with learning difficulties and that it is better to separate control-action by using a joystick for navigation and mouse for interaction selection (Lannen, Brown and Powell, 2000; Lannen, Brown and Standen, 2002; Standen et al., 2004).

A feature of learning disabilities research in ICDVRAT has been user involvement in design of VEs and interaction methods. User-centred design methodology applied in the Virtual City Project (Brown, Kerr and Bayon, 1998) included a user group of 10 students from the Shepherd School in Nottingham who were directly involved in making design decisions during interactive VE development. The entire user group attended the second ECDVRAT conference to present their involvement in the project and their view of this kind of participatory design (Meakin et al., 1998). Subsequent projects have applied this methodology specifically for design development of interaction devices (Lannen, Brown and Standen, 2002; Brown, Shopland and Lewis, 2002; Anderton, Standen and Avory, 2004; Battersby et al., 2004).

3.5 Wheelchair users

Virtual reality and associated technologies have been developed to provide assistive control and training of wheelchair control. Gunderson, Smith and Abbott (1996) presented a concept realisation of a combined human-control/autonomous control system supported by sensor control for collision avoidance. Peussa, Virtanen and Johansson (1998) demonstrated a prototype system using ultrasonic range sensors. Desbonnet, Cox and Rahman (1998) explored use of VE modelling for wheelchair control training. At this stage the application was limited due to low levels of visual realism and crude software modelling. The system was improved and tested with children. These studies found that the nature of disability affected usability of the VE Mobility Simulator (VEMS) (Adelola, Cox and Rahman, 2002) and other tests of virtual environments for training wheelchair control found that manoeuvrability was harder in the virtual environment than in the real world (Harrison et al., 2000). It is concluded that care must be taken ensure that these training simulators are suitable for individual training needs and purpose.

4. APPLICATIONS

4.1 Access and interaction

A considerable amount of ICDVRAT research has examined access and interaction methods – either to provide access to computers for users for whom traditional interface methods are not appropriate, or to develop new interaction methods required by new 3D multimedia environments. Access and interaction methods described include:

- Body movement (camera tracking of user movement) (Madritsch, 1996; Foyle and McCrindle, 2004)
- Mouse emulation: Head controlled mouse emulator for interaction with virtual keyboard (Coyle et al., 1998). Gaze controlled interaction with virtual keyboard (Istance, Spinner and Howarth, 1996). Use of force-feedback mouse with software to detect and compensate for uncontrolled movement such as spasm (Langdon et al., 2002).
- Interaction with virtual agents: navigation using speech, behaviour and gaze (Nijholt et al., 2000).
- Tactile access to 3D graphical information and aural exploration of 3D environments for users with visual impairment (described in §3.1).
- Software adaptation to support multi-modal activity in Windows (Glinert and Wise, 1996; McCrindle and Adams, 1998; Haverty, 2004).
- Use of devices for VE navigation and interaction by people with LD (described in §3.6)
- Successful development of sonar in 3D VE games such as audio space invaders (McCrindle and Symons, 2000) and the Terraformers real-time 3D game with sound interface (Westin, 2004).

4.2 Training and education

Virtual environments for education have been applied to children with intellectual and learning disabilities (described in §3.6) and cognitive development of children with visual impairment (described in §3.1). They have also been applied to evaluation and training of spatial awareness in children with physical disabilities. Stanton, Wilson and Foreman (1996) found 3D training to be better than 2D training for navigational spatial task performance and evidence of transfer of training (of spatial skill) from one VE to another (Stanton et al., 2000) although it could not be determined how much of this was due to features of virtual environment rather than any form of training. These researchers found evidence of vertical asymmetry in spatial memory, in which downward spatial judgements were more accurate than upward spatial judgement, presenting implications for design and use of multi-level VEs for training (Stanton et al., 2002).

A number of research groups have examined use of virtual environments for travel training. A virtual environment for safe street crossing has been tested with stroke patients (Naveh, Katz and Weiss, 2000; Katz et al., 2004) and found that virtual environment intervention was effective in improving measures of visuo-spatial tests and post-VE real world performance in road crossing. A 2D virtual reality street crossing simulation pilot tested with stroke patients found that subjects found it easy to use but requested much more content in the visual scene and additional tasks to do (Lam et al., 2004). An interesting observation from this study was that subjects preferred a 3rd person view of the avatar. Preference for the 3rd person avatar viewpoint was also found in similar research investigating use of 3D virtual environments for travel training of individuals with learning disabilities (Shopland et al., 2004). More recently, virtual environments have been applied to support training of individuals with travel phobia (Sik Lányi et al., 2004).

It has been considered that virtual environments may provide an ideal medium for training for individuals with autistic spectrum disorders (ASD), although consideration must be given to content, design and layout of the virtual environments for this user group (Charitos et al. 2000; Dautenhahn, 2000; Parsons et al., 2000). Later experimental studies found that virtual environments could be used to support training of appropriate social behaviour within one context, although students could not generalise their learning to a different social context (Leonard, Mitchell and Parsons, 2002). This research concluded that virtual environments could successfully be used for education and training but not in isolation. The virtual environment should be regarded as a teaching tool, and is best facilitated by educators (Neale et al., 2002).

4.3 Assessment and rehabilitation

Extensive research has examined use of virtual reality for assessment and rehabilitation of cognitive function. The ImmersaDesk system was used to assess cognitive and functional impairment in patients with traumatic brain injury (TBI), neurological disorder and learning disabilities. Development of a VE-delivered neurological test battery incorporating mental-rotation and reaction time tasks, memory assessment, measures of target acquisition and target recall have been presented by Rizzo et al. (1998, 2000, 2002, 2004). Findings have demonstrated potential as a cost-effective, scaleable tool for attention performance measurement (Rizzo et al., 2002). Others have described development and application of virtual environments for rehabilitation of executive function skills (da Costa et al., 2000; Lo Priore, Castelnuovo and Liccione, 2002).

It has been suggested that virtual reality could be used for postural assessment and vestibular rehabilitation (Alpini et al., 1998; 2000). Use of a computerised system to present bio-feedback of patient centre of pressure in neurological patients found positive, but not conclusive, results with patients with multiple sclerosis (Cattaneo and Cardini, 2000). They concluded that virtual reality could provide a reliable data collection method but needed further developing. A study by Keshner et al. (2004) used projection-based virtual environments, with CrystalEyes shutter glasses to immerse patients in the virtual environment. They measured postural responses to motion of the visual field (sled or scene motion) and found different results for young compared with old adults. Nyberg et al. (2004) found that elderly subjects walked more slowly when immersed in a VE using a HMD and that balance control was most affected by tilting of the visual scene.

Virtual environments have been developed for rehabilitative rehearsal of everyday activities. The Virtual Kitchen provided a number of activities focused around a virtual coffee-making task. Initial pilot studies with Occupational Therapists and patients with traumatic brain injury (TBI) identified design and interface issues (Davies et al., 1998) and recommendations for use of click and drag interaction metaphor and automatic navigation control for users with TBI (Lindén et al., 2000). Case studies of patients using the virtual kitchen and a virtual ATM (automatic teller machine) suggested that virtual reality could be a valuable training tool for patients with TBI, although concern was raised that the development of training environments may not be cost-effective (Davies et al., 2002). Recommendations were proposed for a modular approach to virtual environment construction that would allow for re-use of generic components of a virtual activity (Wallergard et al., 2002).

A virtual kitchen and hot drink-making task was also developed for stroke rehabilitation (Hilton, Cobb and Pridmore, 2000). In an attempt to support rehabilitation of functional performance of this task in addition to cognitive aspects, a tangible user interface was developed allowing direct manipulation of real objects to control the virtual task (Hilton et al., 2002). Lack of flexibility of a tethered interface design led to early stage development of a mixed-reality interface using camera vision to monitor user selection and movement of real objects (Pridmore et al., 2004). However, a clinical pilot study identified limitations of the virtual environment task for rehabilitation of cognitive skills, suggesting that it may *increase* cognitive demand and offer no benefit over supervised real world rehearsal (Edmans et al., 2004).

Greater success appears to have resulted from use of virtual environments for rehabilitation of motor control and coordination of stroke patients. An early demonstration of the potential use of VR technology for rehabilitation of patients with brain injury was the exercise bike used to control navigation (Johnson, Rushton and Shaw, 1996).

Sonic movement environments use auditory feedback as a motivation for patient movement. Tarnanas and Kikis (2002) compared visual feedback with auditory feedback and no feedback for children with learning disabilities. The study found auditory feedback did help with developments in kinaesthesia, motor planning, sequencing and timing capabilities. Lewis-Brooks and Hasselblad (2004) demonstrated potential for use of aesthetic resonant environments as an effective medium providing interactive therapeutic exercises to encourage body awareness, co-ordination and movement in children with physical and cognitive disability. A conceptual model for use of Soundscapes for home-internet based rehabilitation for stroke patients was also presented (Lewis-Brooks, 2004).

4.4 Mobility aids

Virtual reality and associated technologies have been applied to enhancing mobility via the development of mobility aids. Much of this research is technology development and includes: sensors mounted onto wheelchair to facilitate obstacle detection and avoidance (Probert, Lee and Kao, 1996); use of wheelchair-mounted sensors and VR technology for remote control over wheelchair (Gunderson, Smith and Abbott, 1996; Peussa, Virtanen and Johansson, 1998); integration with VR to provide salient information to user (Everingham et al., 1998) and for simulated training of wheelchair use (Desbonnet, Cox and Rahman, 1998; Niniss and Nadif, 2000); intelligent systems providing visually impaired users with information of the hazards around them – the Robotics Travel Aid (RoTA) (Mori and Katani, 1998; Mori et al., 2002) and the Pedestrian Intelligent Transportation System (P-ITS) (Sasaki et al., 2000; Kuroda et al., 2002).

4.5 Language and communication

Virtual reality and associated technologies have also been applied to support of language development and communication for users with speech and/or hearing impairment. Much of this research has been on development and testing of software algorithms to recognise sign language (Losson and Vannobel, 1998; Kuroda et al., 2004) for conversion to speech (Vamplew, 1996) or for animated representation of sign to be used for training people recognise sign language (Tabata et al., 2000; Papadogiorgaki et al., 2004).

However, the work of Sawada's group has shown that technology can also be used to good effect in the areas of improvement of speech synthesis techniques for oesophageal speech (Hisada and Sawada, 2002; Sawada, Takeuchi and Hisada, 2004), whilst others have developed systems for speech therapy: for example Vicsi and Váry (2002) studies in language teaching and training support showed the effectiveness of using visual feedback to help with pronunciation.

4.6 Technology for professional use

Virtual networks have been proposed as an aid for therapists (Magnusson, 1996; Magnusson, 2000; Teittinen and Väätäinen, 2000), collaborative networks between therapists and patients (Almeida and Ramos, 2000; Olsson, Carignan and Tang, 2004), simulation training to enable caregivers to experience stroke (Maxhall et al., 2004) and for medical simulation: knee surgery (Hollands and Trowbridge, 1996) and for surgical training (Al-khalifah and Roberts, 2004).

4.7 Virtual environments as design tools

Early consideration of virtual reality suggested that it may be an ideal medium for evaluation of adaptations made to the home or workplace for disabled users. A prototype system demonstrated potential for integration of computer-aided design models to allow for visualisation of design and user walkthrough of the adapted environment (Davies and Eriksson, 1996). A motion platform integrated with a virtual reality demonstrator system was used to evaluate buildings and environments designed for wheelchair users (Harrison et al., 2000). The HabiTest 3D environment builder demonstrated in 2004 was evaluated by people with physical disabilities with positive results from user testing (Palmon et al., 2004). Virtual environment modelling has also been applied to testing of rehabilitation aids by incorporating anthropometric human models into the VE design (Beitler and Foulds, 1998) or by allowing users themselves to view and evaluate rehabilitation products (Nichols et al., 2002). This latter study found that the VE model was acceptable to elderly patients and could be used to evaluate products, although they did not make use of all of the features of the virtual environment system (zooming, alternative viewpoints, etc.).

5. PROMISE OR PRACTICE – HOW FAR HAVE WE PROGRESSED?

The first ECDVRAT offered suggestions for, and early demonstration of, potential applications of virtual reality and associated technologies for disability. Inter-mingled with expectations for technologies to enhance and improve quality of life, concerns were also raised that we should not throw technology at this community as a 'solution looking for a problem' (Wann, 1996), and that we should ensure that the technology itself did not present users with adverse experiences, including side effects of VR-immersion (Rose, 1996; Zwern and Goodrich, 1996).

Ten years on, we know that these concerns were not barriers to effective implementation of technology for disability. Studies found that side effects were mild and transient in stroke patients (Crosbie et al., 2004) and patients with neurological impairment were at no higher risk of experiencing these symptoms than healthy subjects (Pugnetti et al., 1998). No side effects have been reported in use of non-immersive systems.

Virtual Reality and Associated Technologies are a range of technologies offering interfacing to, and interaction with, multi-media computers, virtual and real environments. They comprise: an 'environment' which can be anything from real world, through networked real world (telecommunication), 'mixed reality' environments, to non-immersive simulated VE, and fully-immersive VE; interaction with computers using devices such as joystick, mouse, 3D controller, gesture and body movement, gaze, and haptic interfaces; feedback from computers via visual, audio, tactile or force-feedback sensory channels; sensor technology such as motion tracking, ultrasound and camera vision; real-time software filtering to compensate for, or enhance feedback to, users with specific interaction and/or sensory requirements.

During this first decade of research, application ideas and potential uses of VRAT have included:

- Rehabilitation following traumatic brain injury (Rose, 1996)
- A means of providing access to computers for visually impaired users (Zwern and Goodrich, 1996)
- A virtual meeting place for representation of emotion (Roberts, Wood and Gibbens, 1996)
- Music therapy (Swingler, 1998)
- Creation of perceptual worlds representing 'inner landscapes' to enhance cognitive learning for visually impaired (Sánchez, Barreiro and Maojo, 2000).
- Interactive painting – movement controlled sound and graphic display (Lewis-Brooks, 2004)
- Rehabilitation that can be offered at home (Loureiro, Collin and Harwin, 2004)

Some of the early promise of this technology is finding its way into practice. Successful application has been demonstrated in use for neurological assessment and in provision of access to computer environments for visually impaired users. Some evidence also indicates realisation of potential for VRAT in rehabilitation of motor control and cognitive skills development.

In many cases, research has not yet reached the stage at which evidence in practice can be demonstrated. In part, this is due to the fact that many projects require years of technical development and pilot testing before they reach experimental testing with target end users. A considerable amount of research effort has contributed to design, development and evaluation of these technologies and some projects have involved large research collaborations between academics, practitioners and users from several different countries.

New areas of application are also emerging, such as home-based rehabilitation systems, pain distraction and exposure therapy. Predominant use of low-cost desktop delivery systems and integration with other technology (for accessibility or assessment of disability) will ensure that the high end and, perhaps, more blue skies approach of some programmes reported through the pages of these Proceedings will find their way into real world solutions for real world problems.

6. CONCLUDING REMARKS

Developing this review of a decade of research and development in disability, virtual reality and associated technologies, raised a number of issues and problems – not least, how to represent in a reasonably concise way the breadth and depth of research in the general area, and also how to structure the contributions. The multidisciplinary nature of the whole research area presents multi-dimensional threads that bind disparate application areas using similar technology or disparate technological solutions applied to the same application area. What has become clear over the past decade is that, as the years have progressed, the ICDVRAT community has developed an impressive body of evidence that virtual reality can and does provide mature alternative solutions. The *esprit de corps* that is so evident in the community bodes well for future collaboration.

Acknowledgements: The authors would like to thank all contributors and sponsors to the growing success of this conference series since its inception in 1996 to the present day.

7. REFERENCES

All papers are from the conference proceedings and can be found online at the conference archive: *www.icdvrat.reading.ac.uk*. For brevity, papers cited in this paper will be referred to below by authorship, paper title and year of publication as follows:

1996: *Proc. 1st European Conference on Disability, Virtual Reality and Associated Technologies*, P M Sharkey (Ed.), Maidenhead, UK, ISBN 070491140X, July 8–10, 1996.

1998: *Proc. 2nd European Conference on Disability, Virtual Reality and Associated Technologies*, P M Sharkey, F D Rose & J-I Lindström (Eds), Mount Billingen, Skövde, Sweden, ISBN 0704911418, Sept. 10–11, 1998.

2000: *Proc. 3rd International Conference on Disability, Virtual Reality and Associated Technologies*, P M Sharkey, A Cesarani, L Pugnetti & A Rizzo (Eds), ISBN 0704911426, Sept. 23–25, 2000.

2002: *Proc. 4th International Conference on Disability, Virtual Reality and Associated Technologies*, P M Sharkey, C Sik Lányi, P J Standen (Eds), Veszprém, Hungary, ISBN 0704911434, Sept. 18–20, 2002.

2004: *Proc. 5th International Conference on Disability, Virtual Reality and Associated Technologies*, P M Sharkey, R J McCrindle, D Brown (Eds), Oxford, UK, ISBN 0704911442, Sept. 20–22, 2004.

I A Adelola, S L Cox and A Rahman, Adaptable virtual reality interface for powered wheelchair training of disabled children, 2002.

A Al-khalifah and D Roberts, Survey of modelling approaches for medical simulators, 2004.

A M Almeida and F Ramos, Collaborative networked framework for the rehabilitation of children with Down's Syndrome, 2000.

D Alpini, A Cesarani, L Pugnetti, L Mendozzi, R Cardini, R Kohen-Raz, A Hahan and G Sambataro, Preventing mobility-related accidents in elderly and disabled, 2000.

D Alpini, L Pugnetti, L Mendozzi, E Barbieri, B Monti and A Cesarani, Virtual reality in vestibular diagnosis and rehabilitation, 1998.

N Anderton, P J Standen and K Avory, Using switch controlled software with people with profound disabilities, 2004.

E A Attree, B M Brooks, F D Rose, T K Andrews, A G Leadbetter and B R Clifford, Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implications for people with disabilities, 1996.

R Bates and H O Istance, Towards eye based virtual environment interaction for users with high-level motor disabilities, 2004.

S J Battersby, D J Brown, P J Standen, N Anderton and M Harrison, Design, development and manufacture of novel assistive and adaptive technology devices, 2004.

M T Beitler and R A Foulds, Anthropometric models of children, 1998.

R Berka and P Slavík, Virtual reality for blind users, 1998.

J Broeren, M Georgsson, M Rydmark and K S Sunnerhagen, Virtual reality in stroke rehabilitation with the assistance of haptics and telemedicine, 2002.

B M Brooks, F D Rose, J Potter, E A Attree, S Jayawardena and A Morling, Assessing stroke patients' ability to remember to perform actions in the future using virtual reality, 2002.

T Brooks, A Camurri, N Canagarajah and S Hasselblad, Interaction with shapes and sounds as a therapy for special needs and rehabilitation, 2002.

D J Brown, S J Kerr and V Bayon, The development of the Virtual City: A user centred approach, D J Brown, S J Kerr and V Bayon, 1998.

D J Brown, N Shopland and J Lewis, Flexible and virtual travel training environments, 2002.

D J Brown and D S Stewart, An emergent methodology for the development of virtual learning environments, 1996.

A Caffrey and R J McCrindle, Developing a multimodal web application, 2004.

D Cattaneo and R Cardini, Computerised system to improve voluntary control of balance in neurological patients, 2000.

D Charitos, G Karadanos, E Sereti, S Triantafillou, S Koukouvinou and D Martakos, Employing virtual reality for aiding the organisation of autistic children behaviour in everyday tasks, 2000.

S V G Cobb, H R Neale and H Reynolds, Evaluation of virtual learning environments, 1998.

C Colwell, H Petrie, D Kornbrot, A Hardwick and S Furner, The use of haptic virtual environments by blind people, 1998.

B B Connor, A M Wing, G W Humphreys, R M Bracewell and D A Harvey, Errorless learning using haptic guidance: research in cognitive rehabilitation following stroke, 2002.

M Cooper and M E Taylor, Ambisonic sound in virtual environments and applications for blind people, 1998.

E D Coyle, M Farrell, R White and B Stewart, Design of a non-contact head-control mouse emulator for use by a quadriplegic operator, 1998.

J J Cromby, P J Standen, J Newman and H Tasker, Successful transfer to the real world of skills practised in a virtual environment by students with severe learning disabilities, 1996.

J H Crosbie, S M McDonough, S Lennon, L Pokluda and M D J McNeill, Virtual reality in the rehabilitation of the upper limb after stroke: the user's perspective, 2004.

R M M da Costa, L A V de Carvalho and D F de Aragon, Virtual city for cognitive rehabilitation, 2000.

K Dautenhahn, Design issues on interactive environments for children with autism, 2000.

R C Davies and J Eriksson, Using virtual reality in the adaptation of environments for disabled people, 1996.

R C Davies, G Johansson, K Boschian, A Lindén, U Minör and B Sonesson, A practical example using virtual reality in the assessment of brain injury, 1998.

R C Davies, E Löfgren, M Wallergård, A Lindén, K Boschian, U Minör, B Sonesson and G Johansson, Three applications of virtual reality for brain injury rehabilitation of daily tasks, 2002.

M Desbonnet, S L Cox and A Rahman, Development and evaluation of a virtual reality based training system for disabled children, 1998.

J A Edmans, J Gladman, M Walker, A Sunderland, A Porter and D Stanton Fraser, Mixed reality environments in stroke rehabilitation: development as rehabilitation tools, 2004.

M R Everingham, B T Thomas, T Troscianko and D Easty, Neural-network virtual reality mobility aid for the severely visually impaired, 1998.

D A Eves and M M Novak, Animated tactile sensations in sensory substitution systems, 1996.

U Feintuch, D Rand, R Kizony and P L Weiss, Promoting research and clinical use of haptic feedback in virtual environments, 2004.

M A Foyle and R J McCrindle, Interaction via motion observation, 2004.

F García-Ugalde, D Gatica-Pérez and V García-Garduño, Segmentation and classification of hand gestures for man-machine communication, 1998.

K Gladstone, H Graupp and C Avizzano, Assessing the utility of dual finger haptic interaction with 3D virtual environments for blind people, 2002.

E P Glinert and G B Wise, Adaptive multi-media interfaces in PolyMestra, 1996.

R W Gunderson, S J Smith and B A Abbott, Applications of virtual reality technology to wheelchair remote steering systems, 1996.

J C Hammond, P M Sharkey and G T Foster, Integrating augmented reality with intelligent home systems, 1996.

A Harrison, G Derwent, A Enticknap, F D Rose and E A Attree, Application of virtual reality technology to the assessment and training of powered wheelchair users, 2000.

C Harrison, P Dall, M Grant, M Granat, T Maver and B Conway, Development of a wheelchair virtual reality platform for use in evaluating wheelchair access, 2000.

R Haverty, New accessibility model for Microsoft windows, 2004.

B Herbelin, P Benzaki, F Riquier, O Renault and D Thalmann, Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy, 2004.

D Hilton, S V G Cobb and T Pridmore, Virtual reality and stroke assessment: therapists perspectives, 2000.

D Hilton, S V G Cobb, T Pridmore and J Gladman, Virtual reality and stroke rehabilitation: a tangible interface to an every day task, 2002.

A Hisada and H Sawada, Real-time clarification of esophageal speech using a comb filter, 2002.

R J Hollands and E A Trowbridge, A virtual reality training tool for the arthroscopic treatment of knee disabilities, 1996.

H O Istance, C Spinner and P A Howarth, Providing motor-impaired users with access to standard graphical user interface (GUI) software via eye-based interaction, 1996.

G Jansson, Can a haptic force feedback display provide visually impaired people with useful information about texture roughness and 3D form of virtual objects?, 1998.

G Jansson, Basic issues concerning visually impaired people's use of haptic displays, 2000.

D A Johnson, S K Rushton and J H Shaw, Virtual reality enriched environments, physical exercise and neuropsychological rehabilitation, 1996.

A W Joyce, III and A C Phalangas, Implementation and capabilities of a virtual interaction system, 1998.

N Katz, H Ring, Y Naveh, R Kizony, U Feintuch and P L Weiss, Interactive virtual environment training for safe street crossing of right hemisphere stroke patients with unilateral spatial neglect, 2004.

D A Keating, The generation of virtual acoustic environments for blind people, 1996.

E A Keshner, R V Kenyon, Y Dhaher and J W Streepey, Employing a virtual environment in postural research and rehabilitation to reveal the impact of visual information, 2004.

R Kizony, N Katz, H Weingarden and P L Weiss, Immersion without encumbrance: adapting a virtual reality system for the rehabilitation of individuals with stroke and spinal cord injury, 2002.

R Kizony, N Katz and P L Weiss, Virtual reality based intervention in rehabilitation: relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke, 2004.

S H Kurniawan, A Sporka, V Nemec and P Slavik, Design and user evaluation of a spatial audio system for blind users, 2004.

T Kuroda, H Sasaki, T Tateishi, K Maeda, Y Yasumuro, Y Manabe and K Chihara, Walking aids based on wearable/ubiquitous computing – aiming at pedestrian's intelligent transport systems, 2002.

T Kuroda, K Sato and K Chihara, S-TEL: An avatar based sign language telecommunication system, 1998.

T Kuroda, Y Tabata, A Goto, H Ikuta and M Murakami, Consumer price data-glove for sign language recognition, 2004.

O Lahav and D Mioduser, Multi-sensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation, and mobility skills, 2000.

O Lahav and D Mioduser, Multi-sensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation, and mobility skills, 2002.

O Lahav and D Mioduser, Blind persons' acquisition of spatial cognitive mapping and orientation skills supported by virtual environment, 2004.

Y S Lam, S F Tam, D W K Man and P L Weiss, Evaluation of a computer-assisted 2D interactive virtual reality system in training street survival skills of people with stroke, 2004.

P M Langdon, F Hwang, S Keates, P J Clarkson and P Robinson, Developing assistive interfaces for motion-impaired users using cursor movement analysis in conjunction with haptic feedback, 2002.

P M Langdon, S Keates, J Clarkson, and P Robinson, Using haptic feedback to enhance computer interaction for motion-impaired users, 2000.

T Lannen, D J Brown and H Powell, Access to virtual learning environments for people with learning difficulties, 2000.

T Lannen, D J Brown and P J Standen, Design of virtual environment input devices for people with moderate to severe learning difficulties – a user-centred approach, 2002.

A Leonard, P Mitchell and S Parsons, Finding a place to sit: a preliminary investigation into the effectiveness of virtual environments for social skills training for people with autistic spectrum disorders, 2002.

A Lewis-Brooks, HUMANICS 1 – a feasibility study to create a home internet based telehealth product to supplement acquired brain injury therapy, 2004.

A Lewis-Brooks, Interactive Painting – an evolving study to facilitate reduced exclusion from classical music concerts for the deaf community, 2004.

A Lewis-Brooks and S Hasselblad, Creating aesthetically resonant environments for the handicapped, elderly and rehabilitation: Sweden, 2004.

A Lindén, R C Davies, K Boschian, U Minör, R Olsson, B Sonesson, M Wallergård and G Johansson, Special considerations for navigation and interaction in virtual environments for people with brain injury, 2000.

C Lo Priore, G Castelnuovo and D Liccione, Virtual environments in cognitive rehabilitation of executive functions, 2002.

O Losson and J-M Vannobel, Sign language formal description and synthesis, 1998.

R C V Loureiro, C F Collin and W S Harwin, Robot aided therapy: challenges ahead for upper limb stroke rehabilitation, 2004.

M Lumbreras, M Barcia and J H Sánchez, A 3D sound hypermedial system for the blind, 1996.

M Lumbreras and J H Sánchez, 3D aural interactive hyperstories for blind children, 1998.

F Madritsch, CCD-camera based optical tracking for human computer interaction, 1996.

M Magnusson, Virtual partners in cyberspace – a proposal, 1996.

M Magnusson, The many rooms of the virtual work-place, 2000.

M Maxhall, A Backman, K Holmlund, L Hedman, B Sondell and G Bucht, Participants responses to a stroke training simulator, 2004.

R J McCrindle and R M Adams, Multimedia Interface for the Disabled (MIND) project, 1998.

R J McCrindle and D Symons, Audio space invaders, 2000.

L Meakin, L Wilkins, C Gent, S Brown, D Moreledge, C Gretton, M Carlisle, C McClean, J Scott, J Constance and A Mallett,User group involvement in the development of a virtual city, 1998.

M Meehan, L Pugnetti, F Riva, E Barbieri, L Mendozzi and E Carmagnani, Peripheral responses to a mental-stress inducing virtual environment experience, 2000.

L Mendozzi, L Pugnetti, E Barbieri, E A Attree, F D Rose, W Moro, A Loconte, B Corrales, L Maj, A Elliot-Square, F Massara and E Cutelli, VIRT – factory trainer project. A generic productive process to train persons with disabilities, 2000.

H Mori and S Kotani, Robotic travel aid for the blind: HARUNOBU–6, 1998.

H Mori, S Kotani, K Saneyoshi, H Sanada,Y Kobayashi and A Mototsune, Research and development project for practical use of robotic travel aid for the visually impaired, 2002.

R G Morris, M Kotitsa, J Bramham, B M Brooks and F D Rose, Virtual reality investigation of strategy formation, rule breaking and prospective memory in patients with focal prefrontal neurosurgical lesions, 2002.

R G Morris, D M Parslow, S Fleminger, B M Brooks, V Giametro and F D Rose, Functional magnetic resonance imaging investigation of allocentric spatial memory using virtual reality in patients with anoxic hippocampal damage, 2002.

R G Morris, D M Parslow and M Recce, Using immersive virtual reality to test allocentric spatial memory impairment following unilateral unilateral temporal lobectomy, 2000.

H J Murphy, Virtual reality and persons with disabilities, 1996.

Y Naveh, N Katz and P L Weiss, The effect of interactive virtual environment training on independent safe street crossing of right CVA patients with unilateral spatial neglect, 2000.

H R Neale, S J Kerr, S V G Cobb and A Leonard, Exploring the role of virtual environments in the special needs classroom, 2002.

S Nichols, P Frost, S V G Cobb, E Marshall, R Eastgate, J H Park, J Armstrong and C D Ward, Virtual environment for assessment of rehabilitation equipment, 2002.

A Nijholt, D Heylen and R Vertegaal, Inhabited interfaces: attentive conversational agents that help, 2000.

H Niniss and A Nadif, Simulation of the behaviour of a powered wheelchair using virtual reality, 2000.

L Nyberg, L Lundin-Olsson, B Sondell, A Backman, K Holmlund, S Eriksson, M Stenvall, E Rosendahl, M Maxhall and G Bucht, Development of a virtual reality system to study tendency of falling among older people, 2004.

A P Olsson, C R Carignan and J Tang, Cooperative control of virtual objects using haptic teleoperation over the internet, 2004.

O Palmon, R Oxman, M Shahar and P L Weiss, Virtual environments as an aid to the design and evaluation of home and work settings for people with physical disabilities, 2004.

M Papadogiorgaki, N Grammalidis, N Sarris and M G Strintzis, Synthesis of virtual reality animation from sign language notation using MPEG-4 body animation parameters, 2004.

S Parsons, L Beardon, H R Neale, G Reynard, R Eastgate, J Wilson, S V G Cobb, S Benford, P Mitchell and E Hopkins, Development of social skills amongst adults with Asperger's Syndrome using virtual environments: the 'AS Interactive' project, 2000.

H Petrie, P Penn, D Kornbrot, S Furner and A Hardwick, Haptic virtual environments for blind people: further explorations with the Phantom device, 2000.

P Peussa, A Virtanen and T Johansson, Improving the mobility of severely disabled, 1998.

T Pridmore, D Hilton, J Green, R Eastgate and S V G Cobb, Mixed reality environments in stroke rehabilitation: interfaces across the real/virtual divide, 2004.

P J Probert, D Lee and G Kao, Interfaces for multi-sensor systems for navigation for the blind, 1996.

L Pugnetti, M Meehan, L Mendozzi, F Riva, E Barbieri, and E Carmagnani, More on central nervous system correlates of virtual reality testing, 2000.

L Pugnetti, L Mendozzi, E Barbieri, A Motta, D Alpini, E A Attree, B M Brooks and F D Rose, Developments of a collaborative research on virtual reality applications for mental health, 1998.

L Pugnetti, L Mendozzi, E Barberi, F D Rose and E A Attree, Nervous system correlates of virtual reality experience, 1996.

D Rand, R Kizony and P L Weiss, Virtual reality rehabilitation for all: Vivid GX versus Sony PlayStation II EyeToy, 2004.

A A Rizzo, T Bowerly, J G Buckwalter, M Schultheis, R Matheis, C Shahabi, U Neumann, L Kim and M Sharifzadeh, Virtual environments for the assessment of attention and memory processes: the virtual classroom and office, 2002.

A A Rizzo, J G Buckwalter, T Bowerly, A van Rooyen, J McGee, C van der Zaag, U Neuman, M Thiebaux, L Kim and C Chua,Virtual environment applications for the assessment and rehabilitation of attention and visuospatial cognitive processes: an update, 2000.

A A Rizzo, J G Buckwalter, P Larson, A van Rooyen, K Kratz, U Neumann, C Kesselman and M Thiebaux, Preliminary findings on a virtual environment targeting human mental rotation/spatial abilities, 1998.

A A Rizzo, L Pryor, R Matheis, M Schultheis, K Ghahremani and A Sey, Memory assessment using graphics-based and panoramic video virtual environments, 2004.

D J Roberts, C Wood and A Gibbens, Tackling isolation and the expression of emotion in a virtual medium, 1996.

F D Rose, Virtual reality in rehabilitation following traumatic brain injury, 1996.

F D Rose, E A Attree, B M Brooks, D M Parslow, P R Penn and N Ambihaipahan, Transfer of training from virtual to real environments, 1998.

F D Rose, B M Brooks and E A Attree, Virtual reality in vocational training of people with learning disabilities, 2000.

Á Sánchez, J M Barreiro and V Maojo, Embodying cognition: a proposal for visualizing mental representations in virtual environments, 2000.

J H Sánchez, AudioBattleShip: blind learners cognition through sound, 2004.

J H Sánchez and H E Flores, AudioMath: blind children learning mathematics through audio, 2004.

J H Sánchez, L Jorquera, E Muñoz and E Valenzuela, VirtualAurea: perception through spatialized sound, 2002.

J H Sánchez and M Lumbreras, Usability and cognitive impact of the interaction with 3D virtual interactive acoustic environments by blind children, 2000.

N Sándor, G Várady, A Szalmás, C Sik Lányi and Z Kálmán, Electronic usage of BLISS symbols, 2002.

H Sasaki, T Tateishi, T Kuroda, Y Manabe and K Chihara, Wearable computer for the blind – aiming at a pedestrians' intelligent transport systems, 2000.

H Sawada, T Notsu and S Hashimoto, Japanese sign-language recognition based on gesture primitives using acceleration sensors and datagloves, 1998.

H Sawada, N Takeuchi and A Hisada, Real-time clarification filter of a dysphonic speech and its evaluation by listening experiments, 2004.

N Shopland, J Lewis, D J Brown and K Dattani-Pitt, Design and evaluation of a flexible travel training environment for use in a supported employment setting, 2004.

C Sik Lányi and Á Molnár-Lányi, Psychological and pedagogic testing of handicapped children with locomotion disorder using multimedia programs, 2000.

C Sik Lanyi, V Simon, L Simon and V Laky, Using virtual public transport for treating phobias, 2004.

M Smyth and J P Wann, Interactive interfaces for movement rehabilitation in virtual environments, 2000.

P J Standen, D J Brown, N Anderton and S Battersby, Problems with control devices experienced by people with intellectual disabilities using virtual environments: a systematic evaluation, 2004.

P J Standen, D J Brown, R Blake and T Proctor, Effective strategies of tutors teaching adults with learning disabilities to use virtual environments, 2000.

P J Standen and H L Low, Do virtual environments promote self directed activity? A study of students with severe learning difficulties learning Makaton sign language, 1996.

D Stanton, N Foreman, P Wilson, H Duffy and R Parnell, Use of virtual environments to acquire spatial understanding of real-world multi-level environments, 2002.

D Stanton, P Wilson and N Foreman, Using virtual reality environments to aid spatial awareness in disabled children, 1996.

D Stanton, P Wilson, N Foreman and H Duffy, Virtual environments as spatial training aids for children and adults with physical disabilities, 2000.

T Swingler, The invisible keyboard in the air: An overview of the educational, therapeutic and creative applications of the EMS Soundbeam, 1998.

Y Tabata, T Kuroda, Y Manabe and K Chihara, Finger character learning system with visual feedback, 2000.

I Tarnanas and V Kikis, Sound-kinetic feedback for virtual therapeutic environments, 2002.

A Teittinen and M Väätäinen, Research and education activities on disability and disabled people at the Virtual University in Nordic countries, 2000.

T ten Kate, P Zizola, B Driessen and K van Woerden, Internet based manipulator telepresence, 2000.

P Vamplew, Recognition of sign language gestures using neural networks, 1996.

K Vicsi and Á Váry, Distinctive training methods and evaluation of a multilingual, multimodal speech training system, 2002.

S A Wall and S Brewster, Providing external memory aids in haptic visualisations for blind computer users, 2004.

M Wallergård, M Cepciansky, A Lindén, R C Davies, K Boschian, U Minör, B Sonesson and G Johansson, Developing virtual vending and automatic service machines for brain injury rehabilitation, 2002.

J P Wann, Virtual reality environments for rehabilitation of perceptual-motor disorders following stroke, 1996.

T Westin, Game accessibility case study: Terraformers – a real-time 3D graphic game, 2004.

F Winberg and S O Hellström, The quest for auditory direct manipulation: the sonified Towers of Hanoi, 2000.

A L Zwern and G L Goodrich, Virtual computer monitor for visually-impaired users, 1996.

ICDVRAT 2006

Session I. Social Interaction

Chair: Tony Brooks

Exploration of social rule violation in patients with focal prefrontal neurosurgical lesions

R G Morris[1], E Pullen[2], S Kerr[3], P R Bullock[4] and R P Selway[5]

[1]Neuropsychology Unit, Department of Psychology, [2]Department of Psychology, Institute of Psychiatry, De Crespigny Park, London, UK

[3] VIRART, Human Factors Research Group, Innovative Technology Research Centre, School M3, University of Nottingham, University Park, Nottingham, UK

[4,5]Department of Neurosurgery, Department of Neurosurgery, King's College Hospital, Denmark Hill, London, UK

[1]*spjtrgm@iop.kcl.ac.uk,* [2]*Emily.Pullen@uclh.nhs.uk,* [3]*Steven.Kerr@nottingham.ac.uk,* [4]*Peter.Bullock@kingsch.nhs.uk,* [5]*Richard.Selway@kingsch.nhs.uk*

[1]*www.iop.kcl.ac.uk*

ABSTRACT

Social rule violation was explored in 22 patients with prefrontal neurosurgical lesions and 22 normal controls. The patients were split into those with neurosurgical lesions impinging on the either the orbitofrontal (OF), dorsolateral (DL) or mesial (M) region of the prefrontal cortex. The study used a virtual reality 'bar' in which participants walked from the entrance to the bar counter, ordered drinks and returned to the entrance, with the choice of moving between other people (socially inappropriate) or around the people (social appropriate). There was a significant increase in socially inappropriate behaviour in the patients whose lesions were in other prefrontal regions than the dorsolateral prefrontal cortex.

1. INTRODUCTION

The ability to follow the social conventions or rules within a society is essential to human social interaction. These rules are learned at an early age through modelling, verbal instruction and reinforcement contingencies. Implicit rules are adhered to in many social settings and when these are violated, this leads to participants feeling uncomfortable and anxious. One particular social convention is to observe the rules concerning personal space and this has been explored using virtual reality. Specifically, Parsons, Mitchell and Leonard (2004) tested the ability of participants to observe social space in a virtual reality bar in which they had to avoid walking between to avatars that were closely facing each other, by taking a longer route to reach the bar. It was found that people with autistic spectrum disorder (ASD) were much more likely walk inappropriately between people than matched controls.

Although a relatively simple measure of social rule violation, this procedure shows that it is sensitive to differences in ASD. It is of interest to establish whether this type of Virtual Reality procedure could also detect social rule violations in other clinical disorders, and in particular those with brain damage. To test, this we studied a group of patients who had undergone focal prefrontal cortical neurosurgery. As well as impairments in the control processes associated with cognition and behaviour (executive functioning) such patients are known to display deficits in the ability to adhere to social conventions and understand social rules, despite in certain cases having previous knowledge of such rules and having experienced normal development (Saver and Damasio, 1991). The prefrontal cortex can be split into heterogenous functional regions, and it has been shown that there are specific associations between the orbitofrontal and mesiofrontal cortex concerning social or emotional function. This includes the demonstration that the damage associated with these regions affects the perception of emotion, the ability learn rapidly changes in social reinforcement, application of rule based behaviour, the modulation of emotional reactions and decision making within a social context (Barrash et al. 2000; Hornak et al. 2005; Morris et al. 2004). Hence, in this study we took patients with focal prefrontal cortical lesions and explored the location of the lesions in relation to putative impairments in the ability to observe social space. It was predicted that orbitofrontal or mesiofrontal rather than dorsolateral lesions would result in such impairment.

 ISBN 07 049 98 65 3

(a) Conversation Blocking / Standing

(b) Conversational Blocking/ Sitting

(c) Social Proximity/ Shopping

(d) Social Proximity/ Standing

Figure 1. *(a) −(d). Starting views of the 4 trials used in the Virtual Reality Task.*

Social impairments in such patients tend to be more subtle than in ASD, so it was necessary to modify the existing Virtual Reality procedure in order to make it more sensitive. This was done by adding in subsidiary tasks to distract the participants from the main requirements of the task, also making it more akin to what might occur in everyday life. The aim was to reduce the amount of planning in relation to the trajectory of participants, such that they were more likely to apply devote less processing resources towards their choice of movement. Additionally, questionnaire measurement of social functioning and executive function was used as comparators in relation to determining the specificity of virtual reality social deficits.

2. METHOD

2.1 Participants

Twenty two patients with prefrontal cortical lesions were included in the study (9 male; 11 female; age: Mean = 40.77; S.D. = 13.34; Education years: mean = 14.27; S.D. = 2.57; National Adult Reading Test predicted IQ: Mean = 107.05; S.D. = 11.70). These patients had no history of psychiatric or major physical illness. They were categorised as to whether they had lesions impinging on the orbitofrontal (OF; n = 14), dorsolateral (DL; n = 13), or mesial (M; n = 14) regions. The patients were compared to 22 normal controls (8 male; 12 female; age: Mean = 40.82; S.D. = 13.50; Education years: mean = 15.59; S.D. = 2.84; National Adult Reading Test predicted IQ: Mean = 109.95; S.D. = 9.19).

2.2 Virtual Reality Task

A virtual bar consists of an approximately square interior with the bar at the far wall opposite the entrance door. In between the entrance door and the bar are two groups of people with a small gap between them. The test consisted of four trials in which the person had to navigate (moving using a joystick) to the bar, order drinks and return the entrance. At the beginning of each trial the participants were therefore faced with a potential direct route to the bar, although the proximity of the two groups of people would make this socially less appropriate. For the four trials, the configuration of the people was varied to depict four different scenes. In the first trial the direct path to the bar was through a gap, either side of which a couple was standing holding a conversation, depicted by their orientation and proximity to one another. In a second trial, there was a similar scene, although the avatars were sitting either side of the gap whilst conversing. These two trials were termed ‘Conversation Blocking.’

The third trial showed a number of shopping bags either side of a gap and the fourth showed two people standing with their backs to one another conversing with a group of people either side of the gap forming the direct path. These two trials were termed ‘Social Proximity.’ Although a direct path did not directly block a conversation between protagonists, our pilot studies of normal adults suggested that it was sufficiently close to the social scenario as to be thought socially inappropriate to pursue the direct path, and the number of direct path trajectories (termed social rule violation) were approximately the same in the ‘Conversational Blocking’ and ‘Social Proximity’ conditions.

A large space existed around the central configuration for participants to walk if they decided not to take the direct path. The remainder of the room consisted of tables occupied by other avatars. There were also two avatars standing to the left of the bar and a 'bar man' standing at the cash till, behind the bar. The task was set up such that the participant moved to the bar to order the drinks and then go back to the entrance, prompting two opportunities for social violation. Sets of simple questions and tasks were developed for participants to carry out on each task and given to participants on a cue card. They could refer to the cue cards throughout the task, thus limiting loading on memory. The tasks comprised of collecting information contained in the bar environment. The aim was for the tasks to provide a distraction from over interpretation of the social task, thus encouraging spontaneous reactions to the social environment and limiting participant bias and response demands.

2.3 Questionnaire and Neuropsychological Assessments

An additional background assessment was conducted to explore the everyday activity and social functioning of the patients using the Patient Competency Rating Scale (PCRS) (Prigitano & Fordyce, 1986), based on an informants report. In addition, the executive functioning of the patients was assessed using the Controlled Oral Word Association Test (COWAT) (Benton, Hamsher & Sivan, 1994); the Hayling Test (Burgess & Shallice, 1997) and the Trail Making Test (Reiten and Wolfson, 1993).

3. ANALYSIS

3.1 Virtual Reality Task

The measure, termed the pathway error, was the number of times the patient took a direct route. Since there were two opportunities for each task for a social rule violation a potential maximum of two could be obtained for each condition. The direct route measures were collapsed for the Conversational Blocking and Social Proximity conditions, producing a mean score (scale 0-4) for each condition, with high scores indicating social inappropriateness.

For the analysis the prefrontal group were split into subgroups in an adaptation of the method described by Rowe et al (2001). In summary, they were divided into those whose lesion impinged on a specific region or otherwise as following: (1) Orbitofrontal cortex (n = 14) versus non-orbitofrontal (n = 8); (2) Mesiofrontal cortex (n = 14) versus non-mesiofrontal (n = 8); and (n = 3) Dorsolateral cortex (n = 13) versus non-dorsolateral (n = 9).

For the virtual reality data, two way ANOVA's were used to explore the data for each comparison with type of Group as between-subject factor (which included the two patient groups created by the division and the control group) and type of task (Conversational Blocking versus Social Proximity) as within-subject factors. Post-hoc analyses were appropriate were conducted to investigate differences between groups within each type of task. For this a more conservative Bonferroni corrected level of $p = 0.03$ or less was used as the threshold of significance.

3.2 Questionnaire and Neuropsychological Assessments

For the questionnaire or neuropsychological assessment data, one-way ANOVA's were used with each variable, and for data that was not normally distributed a non-parametric approach was used with the Kruskal Wallis test to analyses between group differences.

4. RESULTS

4.1 Virtual Reality Task

The pathway errors for the different comparisons are shown in Tables 1. For the analysis including the Orbitofrontal and non-orbitofrontal division there were no main effects of Group or Task. There was an interaction between the factors and this is indicative of the orbitofrontal group showing relatively greater pathway errors on the Social Proximity measure. For the analysis of the mesiofrontal and non-mesiofrontal division, a similar result was found. For the dorsolateral and non-orbitofrontal division there were no Group or Task effects, but there was a significant interaction between Group and Task. A post hoc analysis revealed that this was due to the non-dorsolateral group making being more likely to take a less socially appropriate track on the Social Proximity tasks.

Table 1. *The Pathway Error scores for the Conversational Blocking and Social Proximity Conditions, also showing the F value and levels of significance for the two-way ANOVA's.*

Orbitofrontal versus non-orbitofrontal division	**Control (n=22)**		**Orbitofrontal Group (n=14)**		**Non-orbitofrontal group (n=8)**			**F-Value**	**p-Value**
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Group**	1.13	0.33
Conversational Blocking	1.82	1.62	2.50	1.79	2.13	1.64	**Level**	0.02	0.89
Social Proximity	2.32	1.67	2.79	1.58	1.25	1.75	**Levels X Groups**	3.00	0.06

Mesiofrontal versus non-mesiofrontal division	**Control (n=22)**		**Mesiofrontal Group (n=14)**		**Non-Mesiofrontal group (n=8**			**F-Value**	**p-Value**
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Group**	1.13	0.87
Conversational Blocking	1.82	1.62	2.43	1.60	2.25	1.98	**Level**	0.32	0.58
Social Proximity	2.32	1.67	2.07	1.64	2.50	2.07	**Levels X Groups**	1.58	0.22

Dorsolateral versus non-dorsolateral division	**Control (n=22)**		**Dorsolateral Group (n=13)**		**Non-Dorsolateral group (n=9)**			**F-Value**	**p-Value**
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Group**	1.81	0.18
Conversational Blocking	1.82	1.62	2.15	1.73	2.67	1.73	**Level**	0.53	0.47
Social Proximity	2.32	1.67	1.46	1.61	3.33	1.41	**Levels X Groups**	3.94	0.03

4.2 Questionnaire and Neuropsychological Assessments

The results for the PCRS and neuropsychological tests are show in Table 2, with the three patient divisions. Where the data was not normally distributed ranges are shown and a non-parametric statistical analysis applied. For the PCRS, there were no differences between the groups. For the neuropsychological tests ANOVA's showed significant differences for all types of group comparison, the data indicating impairments in the different patient groups. For the Hayling Test (summary Scaled Score) there were again impairments throughout the patient group. On the Trail Making Test, however, there were no such differences.

Table 2. *Questionnaire and Neuropsychological Assessment Results.*

	Control (n=22)		Orbitofrontal Group (n=14)		Non-Orbitofrontal Group (n=8)		F/ H Value	Sig
	Mean/ Median	S.D./ Range	Mean/ Median	S. D./ Range	Mean/ Median	S. D./ Range		
PCRS Total	140.00	117-150	136.50	101-147	134.50	123-145	0.64	0.43
COWAT	51.32	13.32	38.00	13.36	27.50	16.51	4.91	0.01
Hayling	6.00	5-8	6.00	5-8	6.0	4-8	6.46	0.04
Trails A	31.50	19-70	30.50	16-45	38.50	21-95	3.71	0.33
Trails B	65.5	43-119	67.50	46-120	82.50	67-165	3.71	0.16

	Control (n=22)		Mesiofrontal Group (n=14)		Non-mesiofrontal Group (n=8)		F/ H Value	Sig
	Mean/ Median	S.D./ Range	Mean/ Median	S. D./ Range	Mean/ Median	S. D./ Range		
PCRS Total	140	117-150	136	101-147	134.50	123-145	0.64	0.43
COWAT	51.32	13.32	38.38	13.56	36.71	16.62	4.95	0.01
Hayling	6.00	5-8	6.00	3-8	6.0	4-7	6.22	0.05
Trails A	31.50	19-70	31.00	16-45	34.00	23-94	1.40	0.49
Trails B	65.5	43-119	68.00	50-120	81.00	46-165	2.31	0.32

	Control (n=22)		Dorsolateral Group (n=13)		Non-dorsolateral Group (n=9)		F/ H Value	Sig
	Mean/ Median	S.D./ Range	Mean/ Median	S. D./ Range	Mean/ Median	S. D./ Range		
PCRS Total	140	117-150	134	104-145	142.00	101-147	3.21	0.20
COWAT	51.32	13.32	37.15	13.72	39.00	16.35	4.96	0.01
Hayling	6.00	5-8	6.00	4-8	6.00	3-7	7.14	0.03
Trails A	31.00	19-70	34.00	16-94	31.00	21-40	0.50	0.78
Trails B	65.5	43-119	80.00	56-165	67.00	46-83	4.25	0.12

5. DISCUSSION

The main measure was the number of times the patient took a direct route and since there were two opportunities for each task for a social rule violation a potential maximum of two could be obtained for each condition. The direct route measures were also collapsed for the Conversational Blocking and Social Proximity conditions. A further analysis comparing the orbitofrontal, non orbitofrontal and control conditions showed a trend towards an interaction between condition and group, which suggested greater social rule violation in the orbitofrontal group compared the controls, specifically in relation to the Social Proximity condition. In the analysis comparing the dorsolateral versus non-dorsolateral patients, this pattern was more pronounced but in a reversed manner, suggesting a specific deficit in patients who did not have dorsolateral lesions, in other words those whose lesions were involving orbitofrontal and mesiofrontal regions. Hence, the results are in keeping with the main prediction, that these regions would be associated with social rule violations relating to personal space, in this case the effect occurring for the Social Proximity condition.

A subsidiary question is why the less appropriate behaviour was found only for the Social Proximity condition? One possible explanation for this result is that this is a more subtle test of the ability to avoid social rule violations. Where two people are facing each other in close proximity, for some people, this scenario is less socially ambiguous and it is possible the social consequences of going between facing people is easier to detect. This difference may not be detected in the control participants, where a variety of other factors may determine a decision to go through the direct route, hence the pathway errors are approximately the same. In the patient group, however, those prone to social rule violations will tend to express this in relation to the more ambiguous Social Proximity condition.

The virtual reality task results can be compared to the PCRS to measure everyday activity and social functioning. This questionnaire proved not sufficiently sensitive to detect changes in the patient group, despite these patient having known difficulties in everyday life. It is possible that Virtual Reality techniques

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

may detect social impairments by engaging the patient in social activity, which may prove more valid than relying on a questionnaire technique with informants. In contrast, neuropsychological deficits relating to executive functioning were measurable in the patient group and these suggest impairments in verbal fluency, generativity and response inhibition.

6. CONCLUSION

In conclusion, this is the first study using virtual reality to assess social functioning in patients with acquired brain lesions. It shows that a simple virtual reality procedure that measures judgement of interpersonal space in social situations is sufficiently sensitivity to measure subtle differences in social functioning in patients with frontal lobe neurosurgical lesions. This was in the absence of the ability of informant questionnaires showing such differences. Hence, virtual reality may provide a sensitive method for measuring social deficit in brain damaged patients with subtle social cognitive impairments that cannot be detected using more 'traditional' questionnaire methods. It also supports the notion that social impairment measured in this fashion is more likely to be associated with orbitofrontal or mesiofrontal brain lesions.

Acknowledgements: The authors are grateful to Dr Sue Cobb of the Human Factors Research Group for facilitating this study.

7. REFERENCES

A Benton, AB Sivan and K Hamsher, NR Varney and O Spreen (1994), *Contributions to Neuropsychological Assessment: A clinical Manual 2nd Edition.* Oxford University Press, New York.

J Barrash, D Tranel and SW Anderson (2000), Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Dev. Psychol,* **18**, 355-381.

PW Burgess and T Shallice (1997), T*he Hayling and Brixton Tests*. Thames Valley Test Company, Bury St Edmunds, UK.

J Hornak, J O'Doherty, J Bramham, ET Rolls, RG Morris, PR Bullock and CE Polkey (2004), Reward-related reversal learning after surgical excisions in orbitofrontal and dorsolateral prefrontal cortex in humans*. Jn Cog Neurosc*, **16**(3), 463-478.

S Parsons, P Mitchell and A Leonard (2004), The use and understanding of virtual environments by adolescents with autistic spectrum disorders. *Jn. Aut. Dev. Disord*., **9**(1), 95-117.

RG Morris, J Bramham and A Rowe (2003), Social Cognition following prefrontal cortical lesions. In *The Social Brain* (M. Brune, Ed), John Wiley Press, Chichester.

GP Prigitano and DJ Fordyce (1986), Cognitive dysfunction and psychosocial adjustment after brain injury. In *Neuropsychological Rehabilitation after Brain Injury* (GP Prigatano, DJ Fordyce, HK Zeiner, JR Rouche, M Pepping and BC Woods, Eds), John Hopkins University Press, Baltimore.

RM Reitan and D Wolfson (1993), The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation 2nd Edition. Neuropsychological Press, Tucson, Arizona.

JL Saver and AR Damasio (1991), Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, **29**(12), 1241-1249.

Exploring interpersonal dynamics between adults and motor disabled children within aesthetic resonant environments

P Lopes-dos-Santos[1], L M Teixeira[2], S Silva[3], M Azeredo[4] and M Barbosa[5]

[1,3,4]School of Psychology and Education, University of Porto,
Rua Dr. Manuel Pereira da Silva, 4200-392, Porto, PORTUGAL

[2,5]Research Centre for Science and Technology in Art, Portuguese Catholic University,
Rua Diogo Botelho, 4169-005, Porto, PORTUGAL

[1]*plsantos@fpce.up.pt*, [2]*lteixeira@porto.ucp.pt*, [3]*zanasilva@clix.pt*,
[4]*mvza2002@yahoo.com.br*, [5]*muflyn@graffiti.net*

ABSTRACT

This paper focuses on interpersonal dynamics between the child with disabilities and the adult monitoring his/her performance in Aesthetic Resonance Environments. Drawing upon a social constructivist approach, a framework for human interactivity was checked against empirical data obtained from the exploratory implementation of an environment intending to stimulate body awareness and enhance movement in a group of six children with severe neuromotor disabilities. Results showed that the adult assumed the role of a facilitator, mediating interactions between children and the technological system. The provided social mediation increased quality of movement and improved levels of engagement in the observed group of participants.

1. INTRODUCTION

The application of the concept of 'aesthetic resonance' within the field of special needs and rehabilitation has been recently documented as a therapeutic tool of potential great value (e.g., Brooks and Hasselblad, 2004). The implementation of Aesthetically Resonant Environments is based on a core of technological resources, which captures and transforms human physical movement in events that produce attractive changes in the proximal multi-sensory environment. Therefore, an Aesthetic Resonance Environment is generated by a responsive technology and refers "(...) to a situation where the response to an intent is so immediate and aesthetically pleasing as to make one forget the physical movement (and often effort) involved in the conveying of the intention" (Brooks, Camurri, et al., 2002, p. 205). Such environments are in line with the key rehabilitation principle, which states that therapy must enable patients to improve their often residual capabilities without causing them unnecessary fatigue and frustration (Kizony, Katz, et al., 2004).

Aesthetically Resonant Environments are assumed to enhance movement through "fun" and aesthetic experience, hence implying a shift from task to play. Regarding more traditional views on rehabilitation, such shift is allowed in a context of non-human interactivity where pleasurable experiences result from the interaction between the individual with disabilities and a (non-human) set of integrated digital devices. The formal role of human interactivity in this framework may become increasingly subsidiary as technological investment is made on individualizing the operating system according to the functionalities, preferences, and needs of each user. Centred on optimizing the relationship between the child with disabilities (C) and the system (S), research on Aesthetically Resonant Environments has paid little attention to the interpersonal dimension.

However, interpersonal processes in Aesthetically Resonant Environments are inherent to the presence of an adult (A) monitoring the child's relation with the system. Even when we conceive the system as a stand-alone source of feedback, the intervention of the adult is likely to occur during CS (child-system) interactivity. This intervention may simply show itself as the act of introducing the system to the child, but it may also extend to higher levels of involvement (e.g., demonstrating possibilities of gesture, encouraging responsiveness, improving child's performance through assistance and guidance). As Brooks (2004a) contends, "(…) even notionally single-user activities often happen together with other people and a purely virtual environment tends to cut people off from their surroundings making it difficult for several collocated

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

people to share an experience. (…) For these reasons purely virtual environments may not be desirable in cases where the participants have a disability” (p. 89).

As a general aim, our paper attempts to contribute to the expansion of a dyadic CS concept of Aesthetic Resonant Environments into a triadic CSA (child-system-adult) framework. Drawing upon the social constructivist view (e.g., Wertsch, 1985; Rogoff, 1993, 2003), we propose an approach of the adult as a mediator in the relation between the child with disabilities and the responsive technological system. According to this perspective, child performance and interpersonal processes are not specifiable as independent entities. Taking “action in social context” as the unit of analysis requires a relational interpretation where the participation of the adult in shared activities is assumed to provide the necessary structure or scaffolding for the child to elaborate and transform his or her demonstrated skills (Petersson and Brooks, 2006).

Findings reported in the present paper came from a research project in which we designed and implemented an Aesthetically Resonant Environment intending to stimulate body awareness and improve quality of movement in children with severe motor disabilities. Since the interpersonal dimension appeared as an important organizer of children’s behaviour during the course of sessions, we decided to look closer at such dimension. Therefore, the more specific goal of this study was to investigate child-adult interactivity variables within the designed Aesthetically Resonant Environment. Did such interactivity affect children’s performance? Did social interactive phenomena exert a strong influence on aspects like the quality of gesture or levels of child engagement in the observed sample?

2. METHOD

2.1 Participants

Participants were six children (3 males and 3 females) with severe neuromotor disabilities associated with cognitive and language impairments. Their age ranged from 12 to 13 years old and they were all attending full-time (re)habilitation services in an institution for handicapped children. Inclusion criteria included the capacity to move upper limbs with a reasonable degree of amplitude, the ability to understand simple instructions (verbally or non-verbally) and to perceive contingencies between one’s own gestures and the delivered visual/auditory feedback. Parents signed the informed consent.

2.2 Setup

Using a colour tracking technique, a responsive digital system was designed that translated local gesture into visual and auditory feedback. Each tracked movement simultaneously generated visual feedback projected onto a screen, and controlled the pitch variation of a MIDI instrument. Several templates were created, with different sorts of stimuli (e.g., notes from different scales versus chords versus percussive sounds). The choice of the program (timbre) was left as an option in the interface, as well as the possibility of shifting or lowering the register of each instrument.

The responsive system was implemented using the Max/MSP programming environment for audio processing components associated with the EyesWeb technology (www.eyesweb.org) for visual components. EyesWeb is an open software platform conceived for multimodal analysis and processing of expressive gesture in movement. Developed at the InfoMus Lab of the University of Genoa, the EyesWeb platform consists of a number of integrated hardware and software modules that can be easily interconnected and expanded. The software modules function as a development environment including several libraries of reusable components, which can be assembled to build patches in a visual programming language.

2.3 Procedure and Coding System

Each participant was observed during four sessions that took place on different days at the centre where they were receiving (re)habilitation services. After entering the experimental room, children were comfortably positioned, facing the screen where visual feedback stimuli were projected. Sessions lasted approximately 25 minutes (Mean = 25 min. 30 sec; Range = 18 min. 20 sec. to 32 min. 10 sec.) and were videotaped using simultaneously two cameras placed in different points (one of them –the front camera– captured child’s behaviours, and the other –the back camera– taped screen events).

For analysis purposes, each session was divided in successive 10 seconds intervals and behavioural units were scored for their occurrence once (and only once) per interval regardless of the number of onsets during the interval. Some of the units were recorded on an event base (the observer noted the presence or the absence of events but ignored their duration) and others were scored on a time base (an event was coded within a given 10-sec interval if it could at least be observed during most of the interval length).

All event base units of the developed coding system apply to the child and are briefly described in Table 1. Full operational definitions, containing detailed behavioral descriptors and examples, are available from the first author.

Table 1 Brief definitions of event base observational categories

Categories	Definition
Non-verbal communication	Contextually appropriate use of conventionalized forms of non-verbal signs to express intentions, desires…
Verbal communication	Contextually appropriate use of recognizable words
Positive affect	Positive facial or vocal expressions (e.g., smiles, laughs)
Negative affect	Expressions of negative mood (facial, vocal or gestural)
Avoidance	Behaviour (visual, postural) used to actively avoid contact or interaction with the adult
Spontaneous gesture	Gesture belonging to the child's typical repertoire performed without modelling by the adult
Imitative gesture	Gesture reproducing a gesture performed by the adult
Creative movement	Gesture involving elaboration of behaviour (e.g., combinations, expansions) performed with an apparent creative intent

Concerning time base units, the coding system assessed children's engagement and Child-System-Adult interactivity.

Engagement was broadly defined as the amount of time children spend interacting with the responsive technological system. Taking as reference the level of activity that each child was typically able to demonstrate, four levels of engagement behaviour were considered:

- *Nonengaged* – The child is passive, unresponsive, and ignoring the adult or the visual/auditory information produced by the technological feedback system.
- *Attentional engagement* – The child is not actively interacting with the system but is paying attention to the adult while he or she interacts with the system.
- *Sparsely active engagement* – The child shows intermittent and non-sustained actions as if not fully interested in interacting with the system.
- *Active engagement* – The child actively interacts with the system performing continuously movements and attending to the delivered feedback.

CSA (Child-System-Adult) interactivity was described by the different interaccional states abstracted from the child and/or the adult behaviour. The following categories were defined:

- *Adult as model* – The adult is mainly directive while demonstrating the system to the child. Demonstrative movements may just involve actions performed in front of the child, or may include physical contact (e.g., adult manipulates the child's hand to obtain feedback from the system). Verbal instructions are also coded in this category.
- *Child-Adult as peers in the system* – Both adult and child interact simultaneously with the system by reciprocally imitating or elaborating each other behaviour (e.g., combinations; expansions). Although non-directive, the adult may take some initiatives, providing props, comments on the child's actions or suggestions to extend the child's activity.
- *Child-Adult as peers outside the system* – Child-adult interactions with no direct intermediation of the technological responsive system (e.g., child engages in play or in direct communication with the adult, "ignoring" the system).

- *Child as performer/Adult as public* – The child spontaneously uses the system with the noticeable purpose of being observed and appreciated by the adult. (e.g., before, after or while interacting with the system the child looks at the adult as if waiting for approval or for some other kind of reinforcement).
- *No social interaction* – When no one of the previous described categories is observed (e.g., the child ignores the adult; the adult is controlling the technological device not paying attention to the child).

In observational research, a principal index of usefulness is the reliability of the used coding system, as measured by interobserver agreement. To establish the reliability of obtained scores, two independent observers coded a total of 120 minutes of the tapes (i.e., 720 10 sec intervals). Reliability was determined for event base categories (as a whole), for engagement behaviour and for Child-System-Adult interactivity, and was computed as the total number of agreements divided by the number of agreements plus disagreements. Levels of reliability were 83% for event base categories, 88% for engagement behaviour, and 85% for Child-System-Adult interactivity.

2.3 Data Analysis

Due to variations in the number of observed intervals, results are described in terms of relative frequencies (RFs). RFs were calculated dividing individual frequencies by the total amount of intervals considered in each unit of analysis. To examine the statistical significance of differences we used non-parametric tests. In the employed tests, original scores are changed from continuous to ordinal scales. As usually recommended in these cases, group results for the dependent variables will not be presented through means but described by their median values.

3. RESULTS

3.1 Child-System-Adult Interactivity

RFs medians for the coded interactive states in each of the four sessions are presented in Table 2. As results from the Friedman two-way analysis of variance test revealed, lengths of ‘No Social Interaction’ and of ‘Adult as Model’ states decreased significantly from session to session.

Table 2. *RFs medians for interactive states in the four sessions.*

Interactive States	1rst.	2nd,	3rd.	4th.	Signif.
No social interaction	.218	.139	.117	.065	$p<.001$
Adult as model	.356	.313	.154	.110	$p<.001$
CA as peers in the system	.108	.173	.229	.252	$p<.001$
CA as peers outside the system	.077	.096	.155	.220	$p<.001$
C as performer/A as public	.253	.241	.302	.348	$p<.003$

Inversely, there was a significant trend for the occurrence of the other three interactive states to augment across sessions.

3.2 Child Engagement and Child-System-Adult Interactivity

Considering each interactive state as a separate unit of analysis, we could verify that, when there was no ongoing social interaction between the child and the adult, instances of the ‘nonengaged’ category were clearly predominant (RFs median = .888). Within this interactive state few instances of ‘active engagement’ (RFs median = .066) and of ‘sparsely active engagement’ (RFs median = .035) levels could be observed. The Wilcoxon matched pairs signed ranks test showed that differences between the first (nonengaged) and the other two distributions (active and sparsely active) were significant ($Z = -2.20$; $p< .03$).

Only two levels of engagement occurred during the ‘adult as model’ interactive state: the ‘attentional engagement’ (RFs median = .682) and the ‘nonengaged’ (RFs median = .318). Compared to the second, the first level has appeared with higher incidence in all six participants. Such predominance reached statistical significance, as indicated by a two-tailed sign test ($p< .04$).

Concerning 'Child-Adult as peers in the system', there was a clear prevalence ($p < .04$; two-tailed sign test) of the 'active engagement' level (RFs median = .902) as compared to the 'sparsely active engagement' one (RFs median = .099). During this state no instances of the 'nonengaged' and of the 'attentional engagement' levels were registered. While 'Child as performer/Adult as public' was an ongoing state, 'active engagement' (RFs median =.872) appeared more frequently than the 'sparsely engagement' (RFs median = .149) level did. Because the opposite trend was observed in one of the six participants (RF for active engagement = .423; RF for sparsely active engagement = .577), the difference could not reach statistical significance ($p > .217$; two-tailed sign test). Since we defined engagement as the amount of time children spend interacting with the system, no engagement behaviour was coded during the so-called state 'Child-Adult as peers outside the system' state.

3.3 Quality of Movement and Child-System-Adult Interactivity

Incidence of spontaneous gestures (RFs median = .897) was higher than incidences of imitative gestures (RFs median = .071) and of creative movements (RFs median = .033). In order to explore possible relations between the quality of movement and CSA interactivity, we have considered the total amount of time intervals assigned to the category of 'creative movement'. Proportional occurrences were, then, calculated taking into account the following interactive states: (A) 'no social interaction', (B) 'adult as model', (C) 'child as performer/adult as public', and (D) 'child-adult as peers in the system' (as CS interactivity could not be observed in 'child-adult as peers outside the system' we ignored this state). The same procedure was followed for the categories of 'imitative gesture' and 'spontaneous gesture'.

Instances of 'creative movement' appeared mostly while the child and the adult were playing 'as peers in the system' (see Fig. 1). Compared to each of the other interactive states this incidence was significantly higher in all cases ($p < .04$; two-tailed sign test). Creative movements were also more frequent when the child was acting in front of the 'adult as public' than during the states of 'adult as model' and of 'no social interaction' (in both cases: $p < .04$; two-tailed sign test).

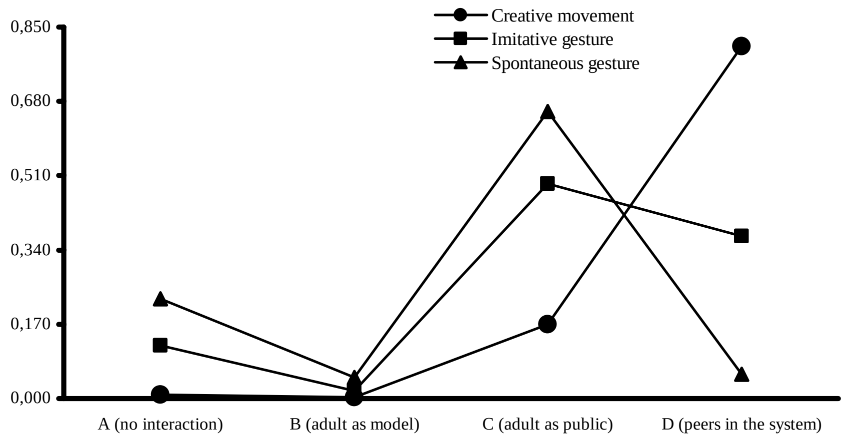


Figure 1. *Medians (RFs) for creative movements, imitative gestures, and spontaneous gestures as function of the considered interactive states.*

Imitative gestures were more frequent within 'child as performer/adult as public' and 'child-adult as peers in the system' than within time segments when 'no social interaction' or 'adult as model' were observed (($p < .04$; two-tailed sign test). Differences concerning spontaneous gestures between the first, the third and the fourth interactive state were not statistically significant.

3.4 Affect Expressions and Child-System-Adult Interactivity

Time intervals assigned to each one of the interactive states were the measures (units of analysis) from which RFs for 'positive affect', 'negative affect', and 'avoidance' have been extracted. Avoidant behaviour was rarely observed. Only three children expressed few instances of avoidance while the adult was acting as a model.

Expressions of negative affect were also seldom observed and merely occurred during the following two interactive states: 'no social interaction' and 'adult as model'.

Concerning the occurrence of positive affect, instances appeared in the states of ‘adult as model’ (RFs median = .290), ‘child as performer/adult as public’ (RFs median = .148), ‘child-adult as peers in the system’ (RFs median = .464), and of ‘child-adult as peers outside the system’ (RFs median = .571).

The Friedman two-way analysis of variance test indicated that differences were overall statistically significant ($p<.006$). Pair wise comparisons by the Wilcoxon test revealed that incidences of positive affect differed between the states of ‘child-adult as peers in the system’ and of ‘adult as model’ ($Z = -1.99$; $p< .05$) as well as between ‘child-adult as peers in the system’ and ‘child as performer/adult as public’ ($Z= -2.20$; $p< .03$). These last two distributions (adult as model and child as performer) differed also significantly from the distribution within the state of ‘child-adult as peers outside the system’ (respectively: $Z= 1.99$; $p< .05$ and $Z= -2.20$; $p< .03$). Differences between events assigned to the states of child-adult as peers in the system and outside the system did not reach statistic significance ($p > .11$).

3.4 Communication and Child-System-Adult Interactivity

Considering the total amount of time intervals assigned to the categories of ‘non-verbal communication’ and of ‘verbal communication’ as the unit of analysis, we explored relations between communicative behaviour and CSA interactivity. One of the participants revealed instances of verbal and non-verbal communication in most of the interactive states. The other participants only demonstrated these behaviours during the states of child-adult as peers (in or outside the system). Both verbal and non-verbal communication behaviours were more frequent in the interactive state of ‘adult-child as peers outside the system’ than in the other referred state (Fig. 2).

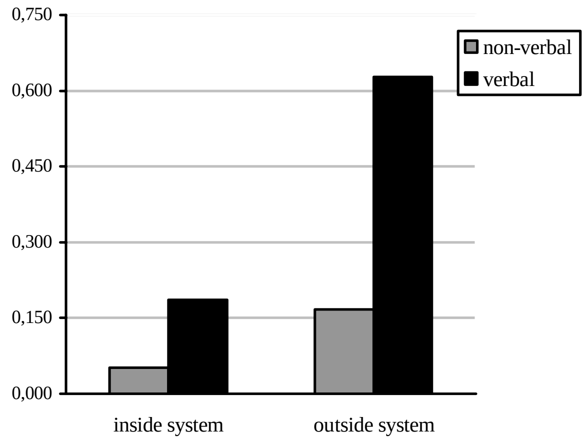


Figure 2. *Medians (RFs) for non-verbal and verbal communication behaviours as function of the considered interactive states.*

Concerning non-verbal and verbal communication, differences between occurrences in both interactive states were significant ($p< .04$; two-tailed sign test).

4. DISCUSSION

The research base for providing (re)habilitation services to children with disabilities has evolved dramatically in the past few years. There is now more complete and detailed information about how humans learn and develop. Contemporary views stress that development occurs across multiple interconnected behavioural levels (Lopes-dos-Santos, 2000; Overton, 2003), and several integrative approaches describe how the social relationships of children with adults and peers provide the critical foundations for learning and development (Spiker, Hebbeler, et al., 2005).

The recognition that social contexts exert a strong influence on child behaviour led us to explore interpersonal dynamics between children with severe neuromotor disabilities and the adult monitoring their performance in a multi-sensorial Aesthetic Resonance Environment. Results showed that adult-child interactions provided an organizing structure for children activities.

Child engagement and expressions of positive affect increased while children and the adult were participating in shared activities, namely when they were acting ‘as peers inside the system’. This interactive state seemed to make visual modelling more efficient, affording the necessary scaffolding for the child to ‘reinvent’ own gestures and to transform them into creative movements. A specific form of modelling – a scenario in which the adult suggested trajectories by moving his/her hand over the screen – appeared as a particularly powerful technique to catalyse subsequent new gestures. Both imitative and creative movements were evoked when the child seemed to act ‘as a performer in front of the adult’. Furthermore, instances of verbal communication were mainly observed when the child and the adult related with each other ‘as peers outside the system’. On the other side, if the adult was demonstrating the system, children were likely to remain quiet and ‘attentively engaged’.

Analysis revealed that joint-activities increased from the first to the fourth observation. There was indeed a trend for instances of ‘child/adult as peers’ (inside or outside the system) to emerge more frequently across sessions. Even though the adult used some facilitation strategies to initiate and maintain the occurrence of such interaccional states, the portrayed results are likely to reflect the effects of practice and learning. Children acquired a progressive capacity to handle the system. Since social behaviour is fuelled by feelings of efficacy following self-recognized achievement of mastery (Shonkoff & Philips, 2000), the enhanced sense of control may have encouraged them to participate in child-adult shared activities.

Specific and new intervention perspectives are revealed when the interpersonal dimension within Aesthetically Resonant Environments is explored. Recognizing the importance of such dimension underscores the potential advantages of integrating social mediation procedures in (re)habilitation frameworks inspired on the concept of ‘aesthetic resonance’.

Disabilities may lead to general developmental problems that become increasingly great, as individuals grow older. Regarding neuromotor-disabled children, evidence shows that most of them do not easily engage in social interactions with others (e.g., Lopes-dos-Santos & Fuertes, 2005). Subsequent delays in the social and communicative domains are, in turn, likely to reduce children’s ability to learn from social environments (Wolery, 2000). The detrimental impact of such disability-amplifying processes can be minimized if interpersonal phenomena that seem to be unique to CSA interactivity are assessed and targeted as intervention goals. Likewise, the same processes that frame social behaviour as a relevant factor in neuromotor (re)habilitation should be seen as possible tools for intervention in other kinds of disabilities – namely in those where social behaviour is itself the primary target (e.g., autistic spectrum disorders).

Practices of embedding interpersonal dynamics in CS interactivity call for more research and further refinements. We are currently conducting an investigation on facilitation strategies to enhance adult-child interactions. Conceptually guided by principles from the ‘Zone of Proximal Development’ construct (Vygotsky, 1978), this work aims at the elaboration of guidelines to improve quality of CSA interactivity within Aesthetically Resonant Environments.

Acknowledgements: Thanks to Carlos Araújo for his expertise in preparing digital versions of the videos. Thanks to the participants and their parents for their unconditional willingness to collaborate in this study. Thanks to the Rehabilitation Center ‘Rui Abrunhosa’ and their staff for the provided collaboration. This work was financially supported by a grant from Secretariado Nacional para a Reabilitação e Integração das Pessoas com Deficiência, Program CITE 2004.

5. REFERENCES

A. L. Brooks, (2004a), Soundscapes, In *Inhabited Information Spaces: Living With Your Data* (D. N. Snowdon, E. F. Churchill & E. Frécon, Eds), Springer-Verlag: London, pp.89-99.

A. L. Brooks, (2004b), Soundscapes: Multi-sensory reciprocity through subliminal (non) control, *Proc. 6th Int. Research Conference: Consciousness Reframed: Qi and Complexity*, Beijing, pp. 37-44.

A. L. Brooks, A. Camurri, N. Canagarajah. and S. Hasselblad (2002), Interaction with shapes and sounds as a therapy for special needs and rehabilitation, *Proc. 4th Int. Conference on Disability, Virtual Reality and Associated Technologies*, Veszprém, pp. 205-212.

A. L. Brooks and S. Hasselblad, S., (2004), Creating aesthetically resonant environments for the handicapped, elderly and rehabilitation: Sweden, *Proc. 5th Int. Conference on Disability, Virtual Reality and Associated Technologies*, Oxford, pp. 191-198.

R. Kizony, N. Katz and P. L. Weiss (2004), Virtual reality based intervention in rehabilitation: Relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke, *Proc. 5th Int. Conference on Disability, Virtual Reality and Associated Technologies*, Oxford, pp. 19-26.

P. Lopes-dos-Santos (2000), The organizational perspective in human development research: A metanarrative framework for guiding transdisciplinary practices in early childhood intervention, *Proc. Int. Conference*: *The Social and Cultural Worlds of the Child*, Braga, pp. 130-138.

P. Lopes-dos-Santos and M. Fuertes (2005), Developmental risk and attachment, In *Development: Family and Educational Contexts* (J. Bairrão, Ed.), Livpsic, Porto, pp.172-199.

W. Overton (2003), Development across the lifespan, In Handbook of Psychology: Vol 6. Developmental Psychology (R. Lerner, M. Easterbrooks, J. Mistry, Vol. Eds.), Wiley, New Jersey, pp. 13-42.

E. Petersson and A. L. Brooks (2006), Virtual and physical toys: Open-ended features for non-formal learning, *CyberPsychol. and Behav.*, *9*, pp. 196-199.

B. Rogoff (1993), Children's guided participation and participatory appropriation in social activity. In, *Development in Context: Acting and Thinking in Specific Environments* (R Wozniak & K. Fisher, Eds.), Erlbaum, New Jersey, pp. 121-143.

B. Rogoff (2003), *The Cultural Nature of Human Development*, Oxford University Press, New York.

J. P. Shonkoff and D. A. Phillips (2000), *From Neurons to Neighborhoods: The Science of Early Childhood Development*, National Academy Press, Washington.

D. Spiker, K. Hebbeler and S. Mallik (2005), Developing and implementing early intervention programs for children with established disabilities, In *The Developmental Systems Approach to Early Intervention* (M. Guralnick, Ed.), Paul Brookes, Baltimore, pp. 305-349.

L. S. Vygotsky (1978), *Mind in Society: The Development of Higher Psychological Processes*, Harvard University Press, Cambridge, MA.

J. V. Wertsch (1985), *Culture, Communication, and Cognition: Vygotskian Perspectives*, Cambridge University Press, New York.

M. Wolery (2000), Behavioral and educational approaches to early intervention, In *Handbook of Early Childhood Intervention* (J. P. Shonkoff & S. M. Meisels, Eds.), Cambridge University Press, New York, pp. 179-203.

Challenges in designing virtual environments training social skills for children with autism

U Andersson, P Josefsson and L Pareto

Laboratory for Interaction Technology, University West
P.O Box 795, SE-45126 Uddevalla, SWEDEN

ulf.andersson@hv.se, pernilla.josefsson@hv.se, lena.pareto@hv.se

ABSTRACT

The purpose of the study is to explore particular challenges faced when designing virtual environments for children with autism, with the purpose of training social skills. Our findings are based on studying autistic behaviour during three years (primary and secondary sources), analysis of related system and other computer mediated assistive technology, as well as general game design. From these studies we have identified eight critical design parameters that need to be adjustable in a system suitable for autistic persons. The parameters importance, their variation range, as well as the need for *independent* adjustment of these were estimated and verified by experienced expert pedagogues.

1. INTRODUCTION

Autism refers to a collection of symptoms, related to disabilities or difficulties in social interaction (Steindal, 1997). There are large variations in capabilities and severity, but the lack of social understanding common in this group often lead to social exclusion (Parsons et al, 2000). Autism including less severe autistic disorders is estimated to affect around 1% of the population in Sweden (Steindal, 1997). Only the minority of these autistic children will manage an independent life as adults. However, social training and increased social competency increase quality of life and reduce the need of care.

General educational software is rarely suitable for this target group, since even though the content might be at an appropriate level, the context in which it appears is often too childish. It must be appropriate for the age of the users, because even though a 15-year old is functioning on the level of a 4-year old, he is still 15. Moreover, many traditional educational games give feedback on erroneous answers, i.e., by running an animation or providing an “error” sound, whereas nothing happens when a correct action is taken. This interaction model is inappropriate for autistic users, since it tend to encourage incorrect behaviour rather than preventing it (Eliasson et al, 1999). Therefore, there is a need for specialized educational software.

Using computer assisted training has become increasingly common to support learners with autism and autistic disorders in their skill improvement. Computer assistance has been used for various purposes, including using robots to enhance social skills (Robins & Dautenhahn, 2004), the use of Excel to improve planning activities (Hart, 2005), a training program for mouse interaction skills (Eliasson et al, 1999), as well as emotion recognition (Blocher, 1999). In particular, Virtual Environments (VE’s) has shown great potential for training social skills (Charitos et al, 2002; Kerr et al, 2002; Kerr, 2002; Leonard, Mitchell, & Parsons, 2002; Parsons et al, 2003; Parsons et al, 2000). Many such systems are based on everyday social situation scenarios, such as going shopping, taking the bus or visiting a café. The main benefit of using VE’s is, according to Parsons et al, (2000), that users can practice skills and social interactions safely, without experiencing potentially dangerous real world consequences and that the scenario can be controlled. Moreover, complexity of the scene can be controlled, and progress pace can be determined by the user. One of the challenges for the VE developer is how to construct the VE to allow freedom of exploration and flexibility in interactive behaviour, without the risk of users missing important learning goals (Kerr et al, 2002). According to (Grynszpan et al., 2005), requirements for software design remains poorly documented in this area.

The aim of this work is to identify critical design parameters, which need to be considered when designing scenario-based VE’s training social skills for users with autistic behaviour. The challenges are due to the heterogeneity of needs and capabilities combined with the intolerance to improper levels of difficulties typical for this target group.

2. AUTISM AND SOCIAL TRAINING

Autism is a collection of symptoms that limits the ability to interact with and understand the surrounding world. These symptoms include reduced social skills, difficulties in understanding contexts and abstractions and a distorted sense of perception. Autistic behaviour often includes a reduced natural interest for human faces and voices (Hadwin & Howlin, 2001). The understanding of other people's emotions and minds is limited (Baron et al., 1988), and both understanding and application of flexible, context dependent rules are problematic (Steindal, 1997). Since they often have problems discriminating between important and unimportant details, care should be taken when designing Virtual Environments (VE) so that important objects and desirable actions can be identified easier. They often have a strong desire to organize their environment around comprehensible routines and regularities. Sudden changes, unpredictable events, unfamiliar environments or demands that are too hard or too easy can appear extremely frustrating, causing undesirable behaviour (Steindal, 1997). They also have a low degree of self-initiative and difficulties in understanding implicit instructions. Implicit choices, i.e., choices without providing alternative responses are generally not appropriate.

Social training is a method for rehabilitation with the purpose of increasing the participant's social skills. Studies show (Liberman & Drake, 1994, Falloon, 1997) that increased social skills have several advantages, including subjective increase in life quality, shorter periods of residential care and longer periods of outpatient care, reduced risk of re-entry to residential care and decreased symptom intensity. One important aspect of social training is to master everyday tasks, like shopping or going to a café.

Routines and behavioural patterns are important for persons with autism, and having to break such habits and rules can cause anxiety and stress. However, part of the social training is to learn to loosen such behavioural patterns in a controlled and cautious way. Finally, the low tolerance for improper levels of challenge and demands means that the VE needs to be adaptable to individual needs and preferences.

Earlier studies show that knowledge acquired in the VE can be transferred into real life situations (Parsons, 2003). Users did however ask for more realistic scenarios since they found the behaviour the computer-controlled characters were too predictable in the VE. Being "just" a simulation, users found it unproblematic to behave in ways inappropriate in real life (for instance traversing private backyards). Therefore, we see a challenge in designing a game-like virtual environment based on realistic social scenarios for training social skills for a wide range of users with autistic behaviour.

3. RESEARCH APPROACH

This work is part of a larger project, aiming at development of a VE training social skills for children and adolescents in the spectrum of high-functioning autism. We have an explorative, user-centric approach to design, and our research team has experiences from interaction design, educational software design, game design, 3D animation and programming, as well as experiences from health and social caretaking of individuals with special needs. One of the team members have worked with autistic children during 4 years. Yet, in the process of trying to understand the needs and requirements of such a system we have realized that the project is far more complex and challenging than we first envisioned. In this paper, our goal is to describe and discuss the particular challenges we have identified to be critical design issues, which are related to special needs of this target group.

We have conducted the following studies: (1) analysed literature concerning autism, related system attempts, other computer mediated assistive technology, as well as general game design; (2) an ethnographic study during three years in a specialized caretaking centre; (3) and 15 in-depth interviews with personnel, teachers and relatives. From this we have derived 8 critical design parameters. Finally, (4) these parameters have been validated by 7 experienced expert pedagogues through estimation tests and interviews.

The ethnographic study took place during four years following 21 autistic children and adolescents, ranging from 6 to 21 years of age, in a caretaking centre. The caretaking centre specializes in children and adolescents with autism and autistic disorders, and they have been in operation since 1981. The centre is unique in the sense that it offers full-time accommodation as well as short-time day care activities. It is one of the most experienced centres in Sweden, and their pedagogical work has attracted attention nation-wide. During this period, we have cautiously observed participants during social interaction and computer usage, trying not to intrude and affect their doing by our presence more than necessary. Being aware of that the presence of strangers can have a strong influence on the behaviour of persons with autistic disorders, these observations was mainly used for an general understanding of relevant issues and as starting points for further explorations.

The first set of interviews took place during 2004 and 2005, and they were all open-ended in-depth interviews of informal qualitative nature. 10 of the interviewees were school personnel, 3 personnel from special accommodations, and 2 were families with autistic children. The age of the involved children ranged from 3 to 21 years. School personnel and accommodation personnel were selected because of their professional experience (2 - 34 years in service with the target group). They had various professional roles, including habilitation assistants, social and special pedagogues. The interviews were aimed at achieving a better understanding of the consequences of autistic symptoms, and their relationship and experience with using technology. The purpose of the study was to acquire understanding and information concerning design issues relevant for a virtual environment training social skills.

To verify the parameters appropriateness and the estimated variation in range, 7 additional interviews were conducted with experienced personnel and teachers, including one member of a national board for evaluating software for children with disabilities. The estimation test was arranged around visual illustrations of the parameters, where the illustrations represented different levels of the proposed parameters. The extreme points were used to create a scale, on which the interviewees were asked to plot the estimated appropriate level for all individuals they knew within the target group. The purpose of the test was to discuss the needed variation span and the importance of the different parameters. The interviews were recorded and analysed and is summarized in this paper.

4. CHALLENGE: VARIATION IN CRITICAL DESIGN PARAMETERS

Autistic people can have a number of different disabilities that are important to take into consideration when designing products adapted to their needs. These include a limited attention span in social situations, distorted sense of perception, difficulties understanding implicit instructions, difficulty discriminating between important and unimportant details, and a low degree of self-initiative. The limited span of attention in social situation influences the ability to recognize and interpret non-verbal communication such as gesture and facial expressions. This is of concern when designing interaction with virtual persons. To compose compound understanding from details can be problematic as well as discriminating between which details are important and which can be ignored.

In (Neale et al., 2002), the importance of flexibility within an educational VE was discussed, regarding allowance of different behaviour in different situations. Forcing different behaviours in successive use of the same situation may also be of interest. Their observations indicate a need to introduce much more randomness within these systems, perhaps by presenting the user with a number of VEs that are (superficially) different. The authors conclude by inquiring systems which allow for a more gradual progression in terms of introducing new concepts, increasing complexity as well as a greater variation in scenarios to prevent habitual actions and ritual behavioural patterns.

Our approach to these problems is to think in terms of a VE with *adjustable parameters*, which can be altered to suit the individual level of each user. Due to the variation in abilities and skills of the user group, and the importance of providing a proper level of challenge in the tasks to be performed, these parameters need to be adjustable as independently of each other as possible. Autistic children do not necessarily follow developmental patterns and can be very skilled in specific areas and yet have great problems with related tasks. Therefore, a model of progression where all lower levels proceed to more advanced levels simultaneously, which is the most common way of making tasks more difficult, is not fine-grained and flexible enough for this user group. The goal is to provide low granularity of change (as also noted by Neale et al, 2002), and a wide range of each parameter, to meet various needs. For this purpose, we have identifying a set of critical design parameters, which can be independently adjustable in a scenario-based virtual environment. These parameters could be adjustable by users, supervisors or intelligent agents in the environment. In the explanation of the parameters below, a shopping scenario is used as illustration.

4.1 Level of variation in scenarios

This parameter represents how much the scenarios change between each session. A low variation of scenarios is equivalent to the scenario looking exactly the same each time it is played, whereas a high variation means that different aspects of the scenario change with each replay. The intended variation is slight transformations of principally the same scenario. Aspects that can vary include, but are not limited to, layout of the scenario, items on the shopping list, number of merchandise in the store as well as their respective location, and different looks and behaviours of the computer-controlled characters. Some aspects of randomness ought to be considered (Neale et al., 2002), which for instance can mean to include additional incidences not in the core path of the scenario, or different formulations of responses to the same action. These kinds of transformations and randomness can be controlled and manipulated as parameters.

4.2 Level of alternative paths through scenarios

Scenarios can have different number of choices along the way towards the goal. The easiest scenario consists of a linear path from start to finish, with no branching or deviation from that path. If the number of choices and alternatives in the scenario increases, the difficulty also increases correspondingly. This parameter does not state whether the different choices are more or less correct in the context, just that more choices are introduced in the scenario. These choices can be either geographical ("Do I go to place A or place B") or more abstract ("Do I pay with money or do I use my Visa?"). A shopping scenario with a low level of alternative paths could be that the user enters a store with only one aisle and one possible path through it, or a shelf with only one kind of apples, whereas a high level can be a store with several parallel aisles or several kinds of apples to choose between.

4.3 Level of difficulty in scenario content

Different and clear levels of difficulty concerning scenario *content* should be considered when designing scenarios. The shop can be small or large, contain few or many merchandise, one or different compartments, one or many cashiers, be empty or crowded to give a few examples. More difficult examples include situations where items on the shopping list are difficult or impossible to find, for instance if the sought items are hidden far away in a corner or the store has ran out of milk for the day.

4.4 Number of steps to complete task

Each scenario contains a number of tasks, e.g. paying at the cash desk. These tasks can either be solved with a single click on a button or icon, or be subdivided into more steps. For instance, when paying for the merchandise, the simplest version just displays a button with the text "Pay for my items" or a picture representing the same. Click the button and you are free to leave the store. The opposite could include clicking a wallet-symbol, drag money-icons from your wallet to the counter representing the correct amount to be paid and then, if the correct amount of money is used, leave the store.

4.5 Level of explicitness of goals

If scenarios are clear with easily describable goals, they are usually manageable for most subjects (Parsons, 2003). Each scenario should have a goal that has to be met. The instructions on how to meet these goals can either be specific or described in general terms. At one end of this parameter, the expectations on the user are very clearly expressed, e.g. "Go in, get an orange, go to the cashier, pay for the orange, and then leave the store". The far end of the range represents more vague instructions such as "Go in and get something tasty to eat". Clear, explicit and apparent goals are recommended for autistic users, and this is an ambition worth striving for. However, in real life this is almost never the case, so such recommendation is in conflict with an ambition of realism. Therefore, this can be seen as a parameter objective to adjustment, where the goal can deliberately be made more "fuzzy" to achieve a more realistic situation and to train the user in acting on such situations.

4.6 Level of richness in details

The difficulty to discriminate important details can be facilitated in at least two ways: by removing details in objects (simplify the appearance of objects) or by removing unimportant objects and details. Both these can be used as parameters by setting attributes describing detail level to objects as well as to visual components of objects, and then only present scenes in the appropriate level of detail. The level of detail in the shopping scene can be altered by increasing the number of items in the store (either total amount or per item), the selection of items to be labelled, or how crowded the store is. Naturally, the level of detail in each object can also vary.

4.7 Level of exploration freedom

According to (Kerr, 2002), actions dependent on free exploration of the VE are not recommended for autistic users. However, since in real life many situations are open for exploration and experimentation where inappropriate actions are possible, this is an aspect that need to be trained nevertheless to become more realistic. Since it is often a difficult task for the users, it must be introduced gradually at an individual pace. This parameter sets the amount of explorative freedom within the scenario. Here we refer to how much the user is geared, explicitly or implicitly, towards a goal. When this parameter is set at its lowest, there is no explorative freedom at all, and the user is only allowed actions leading to progression in the task. In the shopping scenario, this would among other things mean that it's only possible to pick up items that correspond with the items on the shopping list. At the other end of the parameter's range there is total explorative freedom, where the user must by own initiatives explore the environment to discover the goal as well as the possible actions. The former can be accomplished by preventing the user to undertake any inappropriate actions (such as leaving the store without paying). Intermediate levels could be to allow passing by the cashier but that this actions result in explicit or implicit indications of its inappropriateness. This aspect is also discussed in Kerr et al, (2002).

4.8 Level of built-in scaffolding support

Educational VEs need support and supervision by teachers/supervisors according to Kerr (2002) and Parsons (2003). Kerr mean, however, that such supervision can be performed by intelligent agents, and need not necessarily be human. Help and supervision can be provided by the system, as help instructions or as intelligent agents guiding the user to make progress in the task. For instance, if an erroneous action is taken by the user, help can be given by explaining why this was inappropriate, by explaining alternative (better actions), or by showing an alternative (better) action. Scaffolding can be built both within the confines of the VE world and as a meta-component. In the shopping scenario the clerk behind the counter can work as an in-world support. He can compare the items in the shopping cart with the items on the shopping list and check if there is something wrong. He can also help with the process of paying for the items.

4.9 Feasibility of adjustable parameters

In order to control the performance within the scenario and to observe the actions of the user, a control system must be included. We are currently developing and evaluating a two-part system that sends information on user behaviour in the VE world to the decision-making-engine. Information is sent at predetermined checkpoints, representing important steps towards completing the scenario. The checkpoints can also query the game engine regarding what action to take. By consulting the scenario dependent set of rules, the game engine can respond appropriately. In this way, personalized and adjustable critical design parameters can be incorporated in the environment

5. VERIFICATION OF CRITICAL PARAMETERS

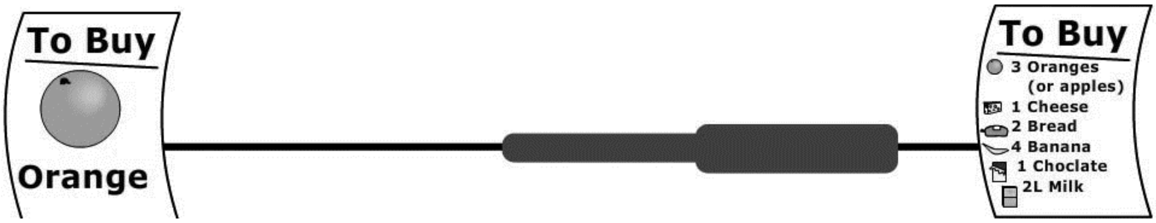
To verify the parameters appropriateness and the estimated variation in range of the different parameters, an estimation test was conducted. To illustrate the parameters and their possible values in an example scenario, pictures of extreme scenes where created and arranged as end points on a scale (see below). The task for the interviewee was to judge the importance of the parameter as such, and to estimate appropriate levels on the scale corresponding to their experience with autistic children. They were asked to give estimations in terms of where on the scale the appropriate level for the lowest performing and the highest performing child they could think of would be, and to give ranges and estimations in percentage where on the scale would be appropriate for other children known to them. The rationale for this procedure was to collect and incorporate all the experience and knowledge the interviewees have with and about autistic children, and all seven interviewees were chosen because of their long experience in the field. The purpose of the test was to discuss the needed variation span, the distribution of abilities and the importance of the different parameters.

All interviewees confirmed that the full spectrum of levels represented by the parameters is definitely needed, as long as it is possible to adjust. They repeatedly expressed the importance of flexibility and adjustment, from the simplest possible to as close to reality as possible. The idea with independent adjustment of the parameters was much appreciated by the software evaluator.

In the diagrams below, the scale and the pictures used in the estimation test is shown. The leftmost picture represents the lowest level in the example scenario, whereas the right most represents the highest. The markings on the line between indicate approximately where the target individuals' needs are located on the scale, according to the estimation test. Two of the parameters where explained without illustrating pictures, and the range was discussed verbally instead.

5.1 Level of variation in scenarios

Variation in scenarios in general is seen as necessary and the possibilities to change variables accordingly are of great importance. To vary the scenario by changing the numbers of products on your shopping list seams to be the most important parameter, while the possibility to change look on the store or the cashier is desirable but not as necessary. The user group often have difficulties with variation since it makes them insecure and unsure of what is expected from them. In a real situation this anxiety can lead to frustration and outbursts. However, since our environment in real life is constantly changing, it is considered very important to practice handling variations. Division in different levels is recommended where the lowest level is with no variations over time, and where the most difficult varies every time. Most autistic children are considered to benefit from training variation in at least one aspect of the scenario, preferably the shopping list, every time the game is played. Variations of the store and its content should be altered by time. The diagram below indicates that the estimation is that most children can mange a shopping list with a few items, if the text is accompanied with illustrative pictures.



5.2 Level of alternative paths through scenarios

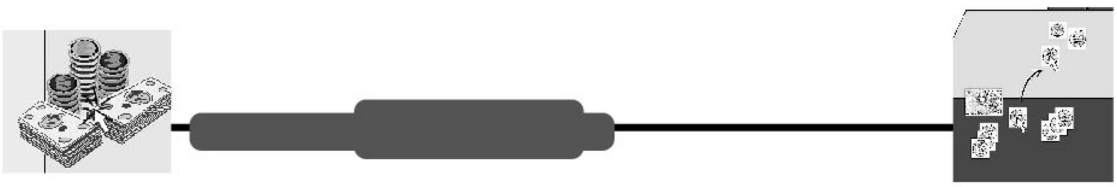
In the left-most scenario, the path is completely determined by the layout of the store, whereas in the right-most the store is large with many possible pathways to choose. The majority of the children (indicated by the markings on the line) needs to be accustomed to an environment where path choices can be made and where the store is possible to investigate.

5.3 Level of difficulty in scenario content

To include various difficulties in the scenarios is considered good, on a more advanced level. Since shops can actually run out of products, this situation is realistic and need to be trained, but the interviewees agreed that this may cause problems for the users. On the other hand, they all agreed that it is good to include these kinds of difficulties, since it can be the starting point for discussions on how to handle unexpected situations in a safe training situation.

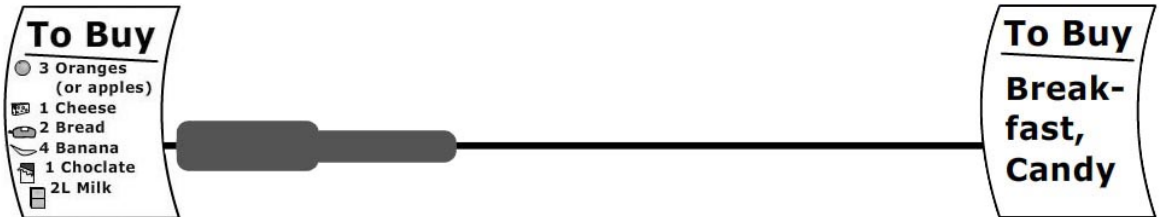
5.4 Number of steps to complete task

Explication should be given to almost everyone because one task will normally involve several steps of acting. A level based explanation is strongly desirable. The possibility to simplify complex tasks by automation of actions is considered an important feature of a virtual environment, since this is something which is more difficult to practice in real life. For instance, the left picture illustrates an automated payment action, where the user only need to click on an icon to get it done. The right picture represents the possibility to leave the exact amount of money asked for by the cashier. However, understanding the value of money is difficult for many autistic children.



5.5 Level of explicitness of goals

In general, an explicit description of the goal needs to be given to most children, to a more or less extent. In the shopping scenario, most children will need to know exactly which products they are expected to buy, one reason being that they need to practice variations in the list and also because making an active choice is often difficult for the autistic child. An implicit goal such as “buy breakfast” involves not only an understanding of the concept of breakfast, but also planning skills, implicit choices and decision making which are difficult.



5.6 Level of richness in details

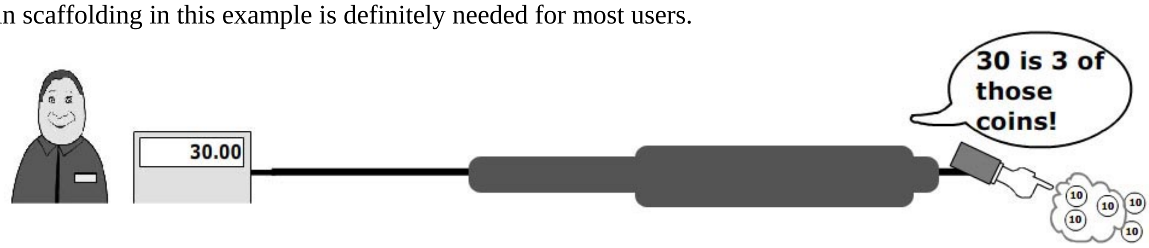
The leftmost picture represents a scenario where the product to be bought in the store is the only that is visible, which means that all objects not involved in the action to be performed is removed from the scene. The rightmost picture illustrates a store with several surrounding items not involved in the action. The majority of the target group copes with that fact that several items exist around them in the store and they will still be able to get the product which was ask for, according to the interviewees. Despite this, they claim that the entire scale from the simplest possible case to a realistic situation is required to cover all various needs. The marking symbolise approximately where most individuals' needs are located.

5.7 Level of exploration freedom

The possibility of exploration freedom, i.e., a world where you can investigate your environment on your own, is considered to retain the player's interests longer. On the other hand, such freedom could be problematic for autistic children, since it requires self-initiatives and self-direction and the ability to concentrate on the task. To mimic reality some exploration freedom is required. For instance, on the higher levels is considered good if an ball can be picked from a shelf, even though it is not on the shopping list (as in real life). It should be possible to kick the ball, but a clear indication of the inappropriate behaviour should be given, preferably accompanied with a suggestion of an accepted alternative way of behaviour.

5.8 Level of built-in scaffolding support

In general, built-in scaffolding support is considered valuable as a complement to human instructors. Software that can be managed by the children without assistance is good for their self esteem and independence. The left-most picture represents no explicit help from the clerk, whereas the right represents an interactive help given by the system. The target group in general have difficulty handling money so built-in scaffolding in this example is definitely needed for most users.



6. CONCLUSIONS AND FUTURE WORK

Our conclusions are as follows:

- The proper level of challenge and difficulty is essential for autistic children, for a successful outcome of their training. Reinforcement of positive behaviour is crucial, so activities should always result in positive emotions and any kind of failure avoided.
- There are no similar games out on the market in Sweden around social training for this particular user group, according to the interviewees. The potential of the suggested VE as a pedagogical tool is considered high, and will likely invoke discussion around social situations among players or adults.

- The estimation test of experienced personnel verify the importance of the critical parameters, and that the suggested variation span is needed
- In contrary to many other level based games, the parameters here need to be *adjustable as independently to each other as possible*, since autistic children can vary extremely in their ability and capacity in different areas, so to adopt to individuals specific needs some parameters may have to be at the lower levels whereas another can be more advanced. Normally, the levels follow each other so that all parameters go from simpler to more advanced levels, which is not fine-grained enough to support this user group.

Future work includes developing the prototype further so that the respective values of the parameters can be tested in a direct manner with autistic users and their companions.

7. REFERENCES

Blocher, Kathi (1999) Affective Social Quests – Teaching emotion recognition with interactive media and wireless expressive toys. MIT, Cambridge

Charitos, D., Karadanos, G., Sereti, E., Triantafillou, S., Koukouvinou, S. & Martakos, D. Employing virtual reality for aiding the organisation of autistic children behaviour in everyday tasks. In P. Sharkey, A Cesarani, L Pugnetti & A Rizzo (Eds) 3rd ICDVRAT, Italy; University of Reading, 147-152, 2000.

Eliasson, E., Fredrikson, A., Rybrand, M., and Wahl, M. 1999. Musbus: a personalized mouse training program for children with autism. In CHI '99 Extended Abstracts on Human Factors in Computing Systems (Pittsburgh, Pennsylvania, May 15 - 20, 1999). CHI '99. ACM Press, New York, NY, 326-327.

Grynszpan, O., Martin, J., and Nadel, J. 2005. Human computer interfaces for autism: assessing the influence of task assignment and output modalities. In CHI '05 Extended Abstracts on Human Factors in Computing Systems (Portland, OR, USA, April 02 - 07, 2005). CHI '05. ACM Press, New York, NY, 1419-1422.

Hadwin, Julie & Howlin, Baron-Cohen (2001). Lära barn med autism att "läsa" andras tankar och känslor. Finland; Cura AB

Hart, M. 2005. Autism/excel study. In Proceedings of the 7th international ACM SIGACCESS Conference on Computers and Accessibility (Baltimore, MD, USA, October 09 - 12, 2005). Assets '05. ACM Press, New York, NY, 136-141.

Kerr, S. J. 2002. Scaffolding: design issues in single & collaborative virtual environments for social skills learning. In Proceedings of the Workshop on Virtual Environments 2002 (Barcelona, Spain, May 30 - 31, 2002). W. Stürzlinger and S. Müller, Eds. ACM International Conference Proceeding Series, vol. 23. Eurographics Association, Aire-la-Ville, Switzerland, 81-91.

Kerr, S. J., Neale, H. R., and Cobb, S. V. 2002. Virtual environments for social skills training: the importance of scaffolding in practice. In Proceedings of the Fifth international ACM Conference on Assistive Technologies (Edinburgh, Scotland, July, 2002). Assets '02. ACM Press, New York, NY, 104-110.

Liberman, P.R. et al. (1986). Social skills training for schizophrenic individuals at risk for relapse, American Journal of Psychiatry 143, 523-526

Leonard, A., Mitchell, P., & Parsons, S. (2002). Finding a place to sit: A preliminary investigation into the effectiveness of virtual environments for social skills training for people with autistic spectrum disorders. In proceedings of International Conference on Disability, Virtual Reality and Associated Technologies, Veszprem, Hungary, University of Reading.

Neale, H.R , Kerr S. J, Cobb, S.V.G, Leonard, A, (2002). Exploring the role of virtual environments in the special needs classroom. In proceedings of International Conference on Disability, Virtual Reality and Associated Technologies, Veszprem, Hungary, University of Reading.

Parsons, Sarah et al (2003) Virtual environments for social skills training: comments from two adolescents with autistic spectrum disorder. University of Nottingham; Nottingham

Parsons, S., Beardon, L., Neale, H.R., Reynard, G., Eastgate, R., Wilson, J.R., Cobb, S.V.G., Benford, S.D., Mitchell, P. and Hopkins, E. (2000). Development of social skills amongst adults with Asperger's Syndrome using virtual environments: the 'AS Interactive' project. In P. Sharkey, A Cesarani, L Pugnetti & A Rizzo (Eds) 3rd ICDVRAT, Sardinia Italy; University of Reading, 163-170.

Robins, B. and Dautenhahn, K. (2004). Interacting with robots: can we encourage social interaction skills in children with autism?. SIGACCESS Access. Comput. , 80 (Sep. 2004), 6-10.

Steindal, Kari (1997) Asperbers syndrom – Att förstå och hjälpa personer med asperbergs syndrom och högfungerande autism. Göteborg; Swedish National association for Autism.

Virtual social environment for preschoolers with autism – preliminary data

C Y Trepagnier, M M Sebrechts, A Finkelmeyer, J Woodford and W Stewart Jr

Department of Psychology, The Catholic University of America,
4001 Harewood Road NE, Washington, D.C. USA

trepagnier@cua.edu

psychology.cua.edu/autism

ABSTRACT

Preliminary results are presented of a feasibility study, still in progress, of a virtual social environment designed to stimulate the social attention of pre-school-aged children with Autism Spectrum Disorder (ASD). The system uses eye-tracking and provides gaze-contingent rewards of clips from preferred videos. Of six children reported on here, most find the experimental setting appealing, and the rewards compelling; they voluntarily engage with it across numerous sessions, and demonstrate learning, with large inter-individual differences in rate of progress. Implications are discussed for the pilot study to follow.

1. INTRODUCTION

Autistic Disorder is characterized by three features, each of which can be present in various degrees: impairments in reciprocal social behavior, impairments in communication, and restricted and/or repetitive patterns of behaviors and interests (American Psychiatric Association, 2000). Autism Spectrum Disorder (ASD) includes Autistic Disorder and related disorders, all of which have in common significant impairments of reciprocal social behavior. A recent study by the Centers for Disease Control and Prevention of the United States has found that 5.5 to 5.7 per thousand children aged 4 to 17 have received a diagnosis of an Autism Spectrum Disorder (ASD) (CDC, 2006).

The value of autism-specific early educational intervention has long been considered incontrovertible (e.g., Harris & Handleman, 2000). However, despite considerable improvements from treatment for some children, there no evidence that any treatment offers a ‘cure’ for this disorder.

ASD is notable for the relative lack of engagement of the affected individual in interactions with other people (e.g., Osterling & Dawson, 1994, Leekam & Ramsden, 2006). Nonverbal communication is severely limited. Young children with ASD may not attend to others’ facial expressions or follow others’ direction of gaze, nor share enjoyment or enthusiasm with others. They attempt to direct others’ attention solely to achieve a goal, e.g., pulling mother’s hand to the doorknob to open the door, and, for some children, pointing to something they want and cannot reach by themselves.

It has been hypothesized that interruption or avoidance of social interaction during critical, experience-expectant early months of life may play a role undermining the child’s further cognitive, social, and communicative development (Trepagnier, 1996). Very young children with ASD, then, may not yet have acquired such basic social skills as following joint attention. The earlier they acquire these skills, the greater their potential to participate in nonverbal interactions of everyday life. Tuning in to nonverbal communication will in turn enable gains in language, social and cognitive development. This hypothesis has motivated the development of a virtual social environment, called Virtual Buddy, to entice very young children with ASD (ages 24 to 60 months) to participate willingly in a training program that draws their attention to the face, and provides multiple experiences of being rewarded for extracting and acting upon information transmitted nonverbally.

 ISBN 07 049 98 65 3

2. STUDY OBJECTIVES

A key feature of the Virtual Buddy system is that it be under the child's control, in order to boost motivation and thereby facilitate learning. Accordingly, training proceeds only when the child chooses to enter the training environment. The primary objective of this feasibility study is to determine the system's acceptability and appeal, over multiple sessions, to a small sample of the varied and challenging population of very young children with ASD, as well as to refine training criteria and other independent parameters. An additional objective is to identify, if possible, individual characteristics associated with progress.

The feasibility study will be completed once data from 12 children participating in up to 20 sessions each have been acquired and analyzed. Results, and on-going observation, are being used to identify needed changes. At the time of writing, 10 children have been enrolled, three of whom are currently in progress.

3. PARTICIPANTS

Results of feasibility trials are reported for six children (one female) ranging in age from 24 to 58 months at the beginning of their participation in the study. These children represented Asian (one), African-American (one), Caucasian (two) and Other (2) ethnicities, reflecting the diversity of the Washington DC area population. Preference was given to children with an older sibling. Four of the six had an older sister, one was an only child, and one had a dizygotic twin sister.

4. PROCEDURE

4.1 Instrumentation

A children's arcade helicopter was modified for the purposes of this project. The interior was gutted and equipped with a tilt-adjustable car safety seat, facing a flat screen monitor and an eye-tracking camera connected to an ISCAN RK-726I system (software version 3.58) in the adjoining room, which also houses the computers and investigator displays. The two rooms are separated by a one-way mirror. The camera and monitor in the helicopter are protected by a Plexiglass shield, and the eye-camera is concealed by darkness. Barriers prevent the child from accessing the side of the helicopter from which cables emerge. The interior of the helicopter is as free as possible of moveable objects that would lend themselves to repetitive play and distract the child from looking at the monitor. The monitor displays bids for attention by a Virtual Buddy, and gaze-contingent video rewards. The video rewards, chosen according to parent report, are individualized to each child.

4.2 Objectives

The objectives of training are to draw the child's attention to the face, and to reward the child not only for looking at the face but also for extracting information from it and other nonverbal behavior, and acting on this information. The specific behaviors addressed by the training that has been implemented are attention to the face and the eye area; and joint-attention-following. Joint-attention-following is operationalized as the child's looking in the direction indicated by the on-screen Virtual Buddy.

4.3 Calibration

The ISCAN eye-tracking system requires calibration, usually accomplished by instructing the participant to look at a particular target and then pressing a key to capture the position of the pupil and the corneal reflection (first Purkinje image) while the participant is looking at that target. The same procedure is carried out for each of 5, or each of 9 targets. Only then can the system acquire eye-tracking data.

To achieve calibration with small children whose language status and cooperativeness may be in question, a semi-automatic calibration procedure is used. Once the child has viewed a couple of minutes of the preferred video, the investigator initiates the capture of one calibration point: the screen shrinks down to a small, 6 by 4 degree area centered on the calibration point, and after 1.5 seconds the eye-tracking program captures the relative position of pupil and corneal reflection for that calibration point. The video then returns to being full-screen. This is repeated, with intervening video watching, until all five calibration points have been captured. Calibration is tested by briefly displaying a bouncing ball with a sound-effect (boing). If the point of regard (POR) veers off from the ball, one or more calibration points can be re-done, while the values for points that are good can be retained.

4.4 Training

4.4.1 Dyadic Attention. For Dyadic Attention (DA) training, a Virtual Buddy, represented by video clips of members of the research team, appears on the monitor, greets the child and offers social praise and comments. Each Buddy appearance constitutes an opportunity to produce the target behavior, gaze at face and eyes. The child's gaze behavior in response to the Buddy receives a composite score reflecting the distribution of gaze among the regions of interest (ROIs): eye area, central face (CF), Buddy and Background. The latency of gaze at the most central region looked at also figures in the score. The components are weighted according to their relative importance, and the score in response to each prompt is compared with a running average of the four previous scores. An improved score earns a reward sequence: the Buddy may say "Wow, you're terrific! Let's watch a video!", and a video ensues. Decreases in score trigger additional cues. If the child is not looking at the face at all, a face-cue is displayed: the screen is masked except for the central face are of the Buddy. If the child has received maximum score for gaze at the CF but not at the Eyes, the screen is masked except for the eyes area.

Dyadic Attention training continues only until the child has demonstrated gazes of .5 seconds at the eye area.

4.4.2 Joint Attention Training. Levels in Joint Attention (JA) training are defined according to the number and salience of the directional cues the Buddy provides. The most salient cuing is provided by manual pointing accompanied by head turn, gaze at the target, and verbalization (e.g., "What's that?"). Initially (Level 1) the response is rewarded as long as the child's gaze moves toward the correct side of the screen (left or right). At Level 2, the child's gaze needs to be detected in the correct ROI. Levels 3 and 4 remove the manual pointing; and Levels 5 and 6 remove the head turn as well, so that only the gaze direction cue is displayed. Figure 1, below, illustrates the three different types of prompt. The first level of each requires only lateral accuracy of gaze.

Figure 1. *Three types of joint attention prompts.*

The reward for looking in the indicated direction is the appearance of a small clip of the child's preferred video, which then expands to fill the whole screen. Four target locations are distinguished, and shown as pastel-colored squares on the screen. Each visible target is located within an invisible ROI, which extends beyond it towards the central window in which the Buddy appears. Detection of the child's POR within the ROI corresponding to the correct target triggers display of the video. If the child's gaze is detected in an incorrect ROI, no video is provided. In case tracking is lost for half or more of the duration of the trial, so that the system cannot detect whether the child looked in the correct direction, a default video is provided. Clips of moving, noise-making toys serve as default videos. Figure 2 depicts three frames captured from the record of a successful response to a manual pointing prompt. In the bottom row of photographs, the change in the child's direction of gaze can be seen by the changing relationship between the two reflections in the child's eye. In the upper row of photographs the cursor representing the child's POR is shown. In the first image the cursor is on the Buddy. In the second it has moved in accordance with the direction indicated by the Buddy. In the third it has arrived at the target. The signal to play the video is sent when gaze is detected in the ROI containing the correct target.

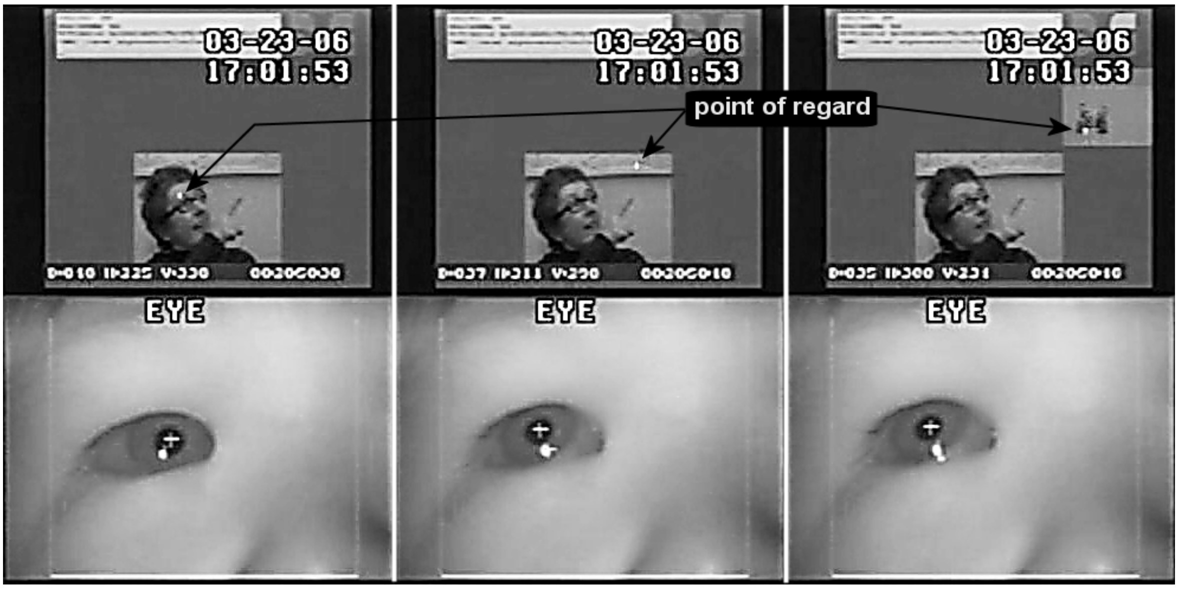


Figure 2. *Following a manual pointing cue to obtain a video reward.*

5. DATA COLLECTION

5.1 Standardized Testing

Clinical diagnosis of an ASD was confirmed by means of the Autism Diagnostic Interview (ADI-R) (Lord et al., 1994), administered by the first author.

Two play-based assessments, the Communication and Symbolic Behavior Scales (CSBS), (Wetherby & Prizant, 1993), and the Mullen Scales of Early Learning (MSEL), (Mullen, 1995), were administered when possible to gain a rounded picture of participating children's developmental and communicative characteristics

5.2 Computerized Acquisition of Gaze Data

Once calibrated, the child's gaze is sampled at 60 Hz. Entry into and exit from the regions of interest defined in DA and JA training are recorded and used in scoring the child's responses, to determine whether a reward is offered and whether the next level of training should begin.

5.3 Probes for Generalization

A series of structured 'real-world' probes was developed, to be offered in the playroom during the 'break'. Probes were designed to be carried out near the beginning of training, mid-way through training, and at the final sessions. The majority of probes were delivered by one of the investigators. These included

5.3.1 Name. The investigator calls the child's name, on entering the room, when the child is facing in the opposite direction, to see whether the child will turn and look;

5.3.2 Request. The investigator tries to elicit requesting behavior by operating some toy which the child enjoys and which requires adult skills, and then waiting to see whether and how the child requests a repetition;

5.3.3 Questioning Intent. The investigator's hand is placed over the child's, while the child is manipulating a toy, blocking the child from continuing to play with it for a few seconds. The objective is to see whether the child looks at the investigator;

5.3.4 Joint Attention. The investigator attracts the child's attention or waits for the child to be looking and then offers one of the bids for joint attention described in Figure 1, to assess joint-attention-following.

5.3.5 Social Praise. The investigator praises the child (e.g., "What a handsome boy you are!") to see whether the child will look and smile.

5.3.6 Sibling Probes. Siblings who were able to participate were coached to carry out two simple probes. The first was to make a teddy bear or other doll carry out a repeated action, like dancing. An identical toy was given to the child with ASD, in order to assess imitation. The second probe was for the sibling to have her doll or teddy bear interact with her sibling's identical toy by hitting or kissing it. Here the objective was to see whether the child with ASD would engage in the game. Some siblings had been videotaped providing the manual pointing prompts, to serve as Buddies.

6. RESULTS

6.1 Acceptability of Training Environment and Schedule

6.1.1 Children's Acceptance of the Training Environment. Most children got into the helicopter and began training at their first session. All but one continued to do so throughout.

6.1.2 Session Length. Sessions lasted approximately one hour. The session always started with the helicopter. After a play break in another room, children usually returned to the helicopter for a shorter, second period. Periods in the helicopter lasted from 4.2 to 33.7 minutes (M = 16.5, SD = 6.5). The helicopter period was ended either by the investigator, or by the child. The child was then escorted to the playroom after the first period in the helicopter, to play with an investigator and participate in probes, or for administration of the play-based assessments. If the child was 'having a bad day' a second period in the helicopter was not offered. If the child insisted, however, a brief second period in the helicopter was offered.

6.1.3 Schedule. Families were invited to come to the lab at as intensive a schedule as they wished (up to 5 times/week). No family chose a 5 days /week schedule. The number of sessions per week ranged from one to four. The majority of families attended two sessions per week.

6.2 Calibration

Calibration was accomplished without any need for children to follow directions, and with minor interruptions to the video. No child showed distress at these interruptions. Because of the stability of children's position, calibration carried out prior to a play break was usable when the child returned to the helicopter in the same session.

6.3 *Training Achieved*

Participants varied in their progress through the training, ranging from no demonstration of learning to completion of all six levels.

6.4 Probes

A longer list of probes was originally planned. The large number of probes resulted in failure to acquire an adequate baseline for all of them because of time limitations. There was no evidence from the probes carried out of carry-over to interaction with investigators or siblings.

6.4 Pre-training Tests

The ADI-R was carried out for 5 of the six children. All met social and communicative criteria for Autistic Disorder, and all had age of onset before three years. None met the repetitive behavior/narrow intense interest criterion. Narrow intense interests are less often seen in pre-school-aged children.

Play-based assessments (the Mullen and the CSBS) were successfully carried out with four of the six children. These data have yet to be analyzed.

7. DISCUSSION

7.1 Child Appeal

Concern that some children might not choose to engage with the training environment was quickly allayed. With the sounds of a preferred, familiar video coming from the helicopter, children were usually eager to get in. An exception was one child who is not among the six reported on here. At the first session his family urged him to get into the helicopter and tried to place him in it against his will. Despite his interest in the video, he had not entered the helicopter by the end of the second session.

Part of the challenge of creating a training environment under the child's control is to make it appealing enough so that the child wants to participate, while at the same time providing rewards contingent on gaze behavior. The continued enthusiasm of five of the six children, the time they spent in the helicopter and their willingness to return to it after the play break demonstrate that the environment is appealing. Data to measure interest in the training, in terms of how often children looked at the screen when the training was going on, have not been extracted. Tracking failure because of looking away from the screen is not distinguished from tracking failure for other reasons, such as a limb in front of the camera or head leaning to one side. However investigators' impression is that after numerous sessions, some children look less often at the training video clips. Because of this observation, ways of reducing the number of sessions and numbers of trials are being considered for the Pilot Study.

Video has proven to be a compelling reward for most children. Often siblings request a turn in the helicopter, and they are given the opportunity while their ASD sibling is in the playroom.

One child used the helicopter for several sessions and then in the midst of the next session became unhappy, and would not reenter the helicopter despite two more visits. Parents are interviewed prior to beginning the study to identify anything the child dislikes or fears (e.g., for one child, barking dogs), so that video clips can be chosen which avoid arousing anxiety. Sometimes, however, the reason for onset of anxiety is not obvious to the observer.

7.2 Session Length and Time in Helicopter

Most sessions begin with the child's entry into the helicopter. As long as the child appears comfortable and engaged, he is allowed to remain in the helicopter at the investigator's discretion. In order to maintain interest and to avoid having young children sitting still for long periods without moving around, periods in the helicopter were usually limited by the investigator, especially with the younger children. To help the child accept leaving the helicopter, video curtains close, the sound is shut off, and the child is invited to come and play with toys in the playroom. Playroom sessions are usually terminated by asking whether the child would like to go back to the helicopter. Sometimes playroom sessions are cut short because the child requests return to the helicopter.

7.3 Schedule

All of the families of the children reported on here, and all but one of the families enrolled in the study so far, live outside of the District of Columbia itself, so that they are driving into the city from counties in Maryland and Virginia, in a traffic environment known for its congestion. Constraints due to distance, traffic, and the child's, sibling's and parent's schedule resulted in the need to hold regular Saturday sessions for two families. On only one occasion did a family fail to show up for a scheduled session without prior email or telephone notification.

7.4 Calibration

The use of a car seat was helpful in that it was familiar, and children were comfortable with having the safety harness attached, or attaching it themselves. The car seat was also slightly tilted back. Both these features helped to maintain the child in a stable position favorable to eye-tracking, and meant that the child returned to the same position, so that calibration did not need to be repeated. There were nevertheless some difficulties with eye-tracking. Data was lost if the child's limb was blocking the camera, whether because legs were extended to the Plexiglas covering the monitor, or because the child's arm was moved in front of the camera. Tracking was lost when the child was leaning to one side while watching the screen. While a wearable tracker might have reduced these losses, mounting of equipment on the child was avoided so that the child would feel free to come and go. In addition, it is likely that pre-school-aged children with ASD would not tolerate a head-mounted display or eye-camera, especially for multiple sessions, and even more likely that they would use it for repetitive play.

7.5 Training

The child who took a dislike to the helicopter after a few sessions was the youngest child in the study (24 months of age at his first session), and did not, in the few sessions in which he took part, demonstrate learning. Two children completed all 6 levels of JA training with few or no errors in the early levels. Their performance suggests that these skills were already established or emerging. These two children were the most advanced in language, with functional, spontaneous phrase speech. Two children demonstrated learning and progressed to Level 6. These two children used speech functionally and spontaneously, but as single-word utterances rather than multi-word phrases. The two children who made little or no progress did not use

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

speech spontaneously. Data from the play assessments will be examined once all twelve children have completed the study, in order to determine whether to use language levels for inclusion and exclusion criteria for participation in the Pilot Study.

7.6 *Generalization*

Probes used were reduced in number, over the course of these children's participation. Some of the probes retained were not carried out early and often enough to be useful in detecting change.

Parents whose children are in the study have the opportunity to monitor their child's performance and see where he is looking in response to the joint attention prompts. The investigator's display includes the child's face and expression as well as the location of the POR superimposed on the video. Some of the parents were interested in watching, while others used the opportunity to devote some one-to-one time to the child's sibling. Although parents were not instructed to transfer the skills being taught to the home, some did so. If this technique were in clinical use, it would be important to include a parent component.

8. CONCLUSIONS

The use of an inviting setting and clips from preferred video has been successful in enticing children to participate in this training program. For the most part children perceive it as a treat rather than an imposition.

Further, there is evidence, for two of the six children, that they are acquiring joint-attention-following skills, which they did not already possess. Training needs to be revised, however, to minimize the numbers of trials, so that children's interest in the training, as opposed to the video, will still be high when the more subtle indicators of direction of attention are being presented. The observations and data acquired from this feasibility study will inform changes to be made prior to carrying out a Pilot Study of efficacy.

Virtual Buddy represents a first attempt to address social skills in this population by means of a potentially stand-alone virtual environment, and the first attempt to provide an intervention that would be freely chosen by children rather than being imposed upon them. Results to date support the feasibility of this approach.

Acknowledgements: We wish to acknowledge the support of the National Institute of Mental Health (MH064924-01A2), and the volunteer families and individuals with ASD without whose participation none of this work would be possible.

9. REFERENCES

S L Harris and J S Handleman (2000), Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *J Autism Dev Disord*, **30**, *2*, pp. 137-142.

S R Leekam and C A Ramsden (2006), Dyadic orienting and joint attention in preschool children with autism. *J Autism Dev Disord*, **36**, *2*, pp. 185-97.

C Lord, M Rutter and A LeCouteur (1994), Autism Diagnostic Interview – Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*, **24**, pp. 659-85.

J Osterling, and G Dawson (1994), Early recognition of children with autism: a study of first birthday home videotapes. *J Autism Dev Disord*, **24**, *3*, pp. 247-58.

C Y Trepagnier (1996), A possible origin for the social and communicative deficits of autism. *Focus on Autism and Other Developmental Disabilities*, **11**, *3*, pp. 170-182.

A M Wetherby and B M Prizant (1993), *Communication and Symbolic Behavior Scales – Manual, Riverside, Chicago.*

ICDVRAT 2006

Session II. Motion Tracking and 3D Modelling

Chair: Cecília Sik Lányi

Development of vision based meeting support system for hearing impaired

R Shikata[1], T Kuroda[2], Y Tabata[3], Y Manabe[4] and K Chihara[4]

[2]Department of Medical Informatics, Kyoto University Hospital,
Kawaramachi, Shogoinn, Sakyo, Kyoto, JAPAN

[3]Department of Radiation Technology, Kyoto College of Medical Technology,
Sonobe, Nantan, Kyoto, JAPAN

[4]Graduate School of Information Science, Nara Institute of Science and Technology,
Takayama, Ikoma, Nara, JAPAN

[1]*ryo@7501.net*, [2]*Tomohiro.Kuroda@kuhp.kyoto-u.ac.jp* , [3]*yoshi-t@kyoto.medtech.ac.jp*,
[4]*manabe@is.naist.jp*, [4]*chihara@is.naist.jp*

[2]*www.kuhp.kyoto-u.ac.jp*, [3]*www.kyoto.medtech.ac.jp*, [4]*chihara.naist.jp*

ABSTRACT

This paper describes a new meeting support system that helps the hearing impaired to understand the contents of the meeting. The proposed system distinguishes the mainstream of the discussion from other chattering based on the utterances. The situation of the meeting is acquired as a picture using an omni-directional vision sensor, and the system analyzes speaker's relations from the captured image by using face directions for the participants. The system shows the mainstream and the chattering of a meeting by using the analyzed result and speech-recognition.

1. INTRODUCTION

In a meeting, the hearing impaired would have difficulty in catching the meeting's contents without note-taker or sign interpreters. The participants in the meeting have some opinions and discuss with another participants. When a participant speaks the opinion, hearing people find the speaker from direction of voice, but the hearing impaired are difficult. The hearing impaired catches the meaning of a speech by lip-reading, when they discuss with a hearing person. The lip-reading is an effective way while they talks to hearing person face to face. Therefore, the hearing impaired has trouble to finding the speaker at the meeting of a large attendance and has trouble to understanding the contents.

There are two ways to let the hearing impaired know contents of meeting. One is to help by sign interpreters or note-takers. The other is to help by using computer system. However, the former has problems that it takes a long time to train sign interpreters or note-takers, and that the shortage of the interpreters occurs. As the latter, the basic technologies like speech recognition have been developed and have been utilized in some kinds of fields. To make the meeting support system, the speech recognition is an effective technology, but most speech recognition method is utilized against a speaker. Thus, it is necessary to have an idea to use the developed speech recognition in the meeting of a large attendance.

This paper proposed a new support system for the hearing impaired. The proposed system captures the scene of meeting and voice data of participants from microphones and omni-directional vision sensor at the same time. The proposed system analyzes the relations of participants from capturer images and translates voice data to character data with speech-recognition technique. The proposed system classifies the translated character data into main speech and chattering by utilizing analyzed data, and displays character data of the main speech.

2. CONCEPTUAL DESIGN OF PROPOSED SYSTEM

2.1 Overview of proposed system

Our proposed meeting support system consists of four parts; image recognition part, speech recognition part, information integration part and display part. Figure.1 shows the conceptual design of proposed system.

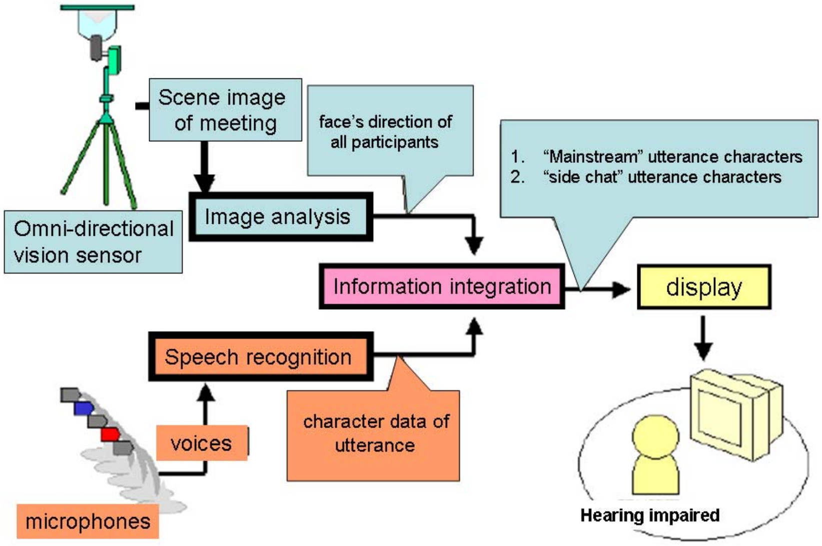


Figure 1. *Conceptual design of proposed system.*

First of all, the proposed system captures scene images of meeting from omni-directional vision sensor and captures voice data of all speakers from microphones. Secondly, the image analysis part in proposed system detects the face directions of all speakers from captured images and analyzes the group of speaking the mainstream of the discussion in the meeting. The speech recognition part translates captured voices to characters. Thirdly, the information integration part distinguishes characters of the "mainstream" and ones of "chat" by using the information from image analysis part and speech recognition part. Lastly, display part shows the mainstream of the discussion as characters. The proposed system would enable the hearing impaired to understand the keystone of the meeting they join.

2.2 Image processing of face direction detection

The mainstream of the discussion would be a set of important utterances in the meeting. Therefore, a speaker, who told an opinion related the mainstream, is the attention at the meeting and other participants of the meeting would watch the speaker. Thus, the detection of face directions enables the proposed system to find the mainstream of meeting. The image recognition part recognizes face directions of meeting's participants from captured omni-directional images.

The proposed system utilizes the omni-directional vision sensor. This vision sensor can capture 360 degree's images. Comparing with using some cameras, the omni-directional vision sensor is not necessary to synchronize the cameras for detecting face directions. This vision sensor is put on the center of table in meeting space to capture face data of all participants. The proposed system estimates face directions in the following manners. First, the system makes a pseudo-panorama image from a captured omni-directional image (Figure.2). Second, the system extracts face areas from the panorama image. Third, the system detects the face direction by using two methods; face recognition with sobel filter and face recognition with color information (R Brunell et all 1993, Y.Dai et all 1995).

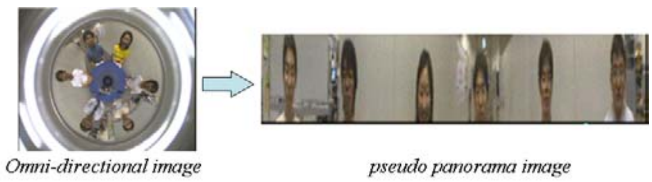


Figure 2. *Omni-directional image and panorama image.*

2.2.1 Extraction of face areas. The system makes one forth resolution image from the pseudo-panorama image from mosaic image process in order to run in a real-time and extracts the skin-color-areas and hair-color-areas from the pseudo-panorama image. Figure3 shows the process of face area extraction.

Proc. 6th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Esbjerg, Denmark, 2006
©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

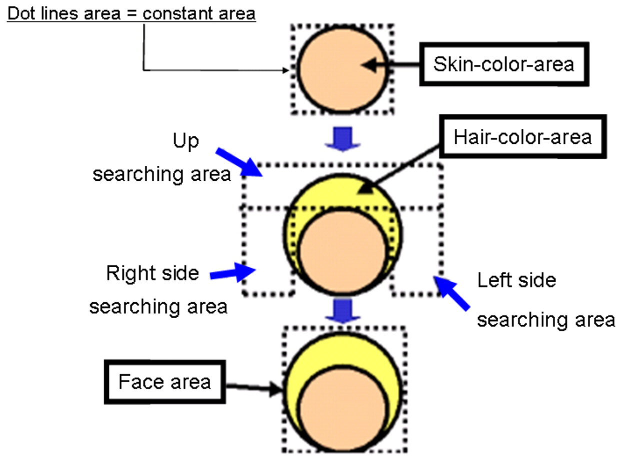


Figure 3. *Overview of face area extraction.*

To extract the skin-color area, firstly, each element in RGB colors is normalized. Next, when the normalized color is in the range of defined color parameters, the system defines the normalized colors as skin-color region. Thirdly, the skin-color regions are labeled. The system defines the labeled as the skin-color area, if the labeled has the constant area. To extract of the hair-color-area, RGB color is converted to Y/Cr/Cb color. If Y/Cr/Cb color is in the range of defined values, the Y/Cr/Cb is defined as the hair-color-area.

The system detects face area by using these color-areas; skin-color-areas and hair-color-areas. The system searches for the hair-color-areas near the skin-color-areas. If the hair area was found, the system defines both the hair-color-areas and skin-color-areas as the face areas.

2.2.2 Detection of face direction. As the method to detect face direction, the system utilizes two detection results of edge information and color information. Edge detection filter, Sobel Filter, is utilized to acquire the edge information. Sobel filter is one of the edge detection filter. When face image is processed by this filter, the outlines of mouth, nostril and eyes in one's face, a large gap of pixel values, are detected.

The proposed system makes a histogram by the acquired edge image. The histogram's horizontal axis is position and its vertical axis is a number of pixels. The histogram has the following features. When a speaker turns toward the front, the outlines of mouth, eyes are centered in the face area. Thus, the edge distribution's median is centered in the histogram. When a speaker turns to the right side, the outlines of face parts are right-sided in face area and the distributions median is also right side. When a speaker turns to left side, the edge distributions median is also left-side. Therefore, the proposed system detects the face direction, front, right and left from using the edge distribution median.

The proposed system calculates the median-value of edge histogram and distinguishes three directions of participant's face. In addition, the proposed system makes the histogram of skin-color area, and it also calculates the median value of the skin-color-histogram. The system recognizes the face direction by using the relation of these median values Figure.4 shows the processing of face direction detection.

If it is difficult to detect the face direction by using the differences between the median value of edge and one of skin-color, the system uses the following processing to detect face direction. The processing is to compare the median value of skin-color-area with the median value of hair-color-area. Both median value of skin color and median value of hair color depend on the face direction. Therefore, it can detect one's face direction to compare these median values.

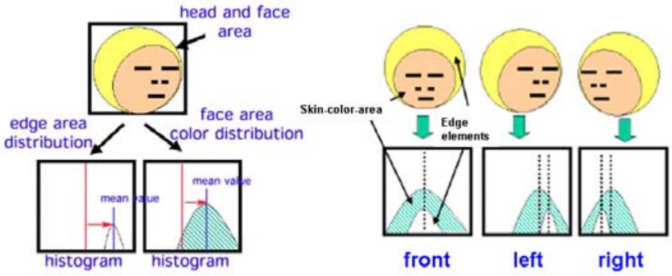


Figure 4. *Face direction detection.*

2.3 Speech recognition Process

It is necessary to display the contents of meeting so as to let the hearing impaired join in the meeting. The speech recognition part translates voice data to characters automatically. Output data from the speech recognition part is character data corresponded to speakers' voices. A number of microphones is equal to one of participants in meeting. To make the prototype system, IBM ViaVoice software is utilized in this paper.

2.4 Information Integration Process

The information integration part, one of the parts in the component of proposed system, classified the speeches of meeting into two groups; a group of "mainstream" and a group of "chat". This part performs the analysis of grouping from both output data of image recognition part and one of speech recognition part

The following procedures are processed in this part. Firstly, it is determined the groups the speaker belongs to from the relation of next participants.Figure.5 shows the relation of next participants to make the groups.

The rules to make groups are shown the below.

[RULE 1] When the proposed system recognized that a speaker and next speaker faced each other from the face direction, they belong to the same group.

[RULE 2] When the proposed system recognized that a speaker and next speaker looked in the opposite direction, they do not belong to the same group, but belong to the difference group each other.

[RULE 3] When it is not included in the RULE1 and RULE2, they do not have the group to belong.

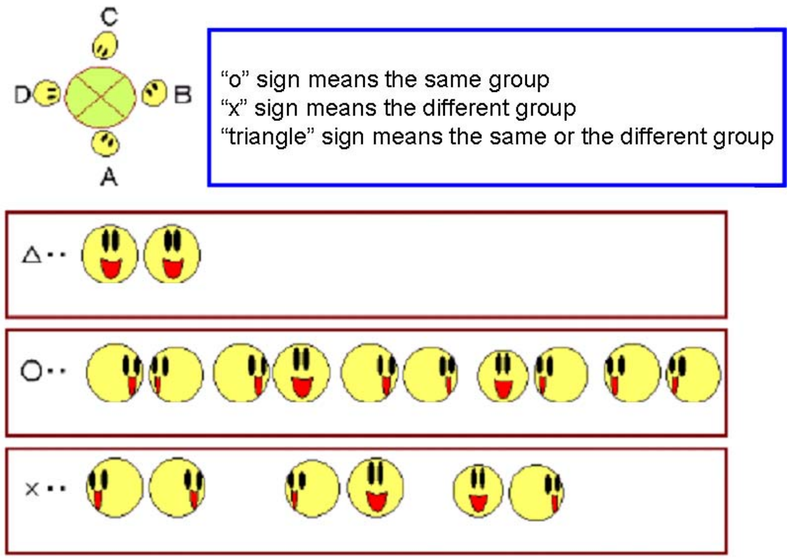


Figure 5. *Rule to make a group between next speakers.*

Secondly, this part makes two groups of "mainstream" and "chat" from groups made under the above rules.

The boundary of a difference group determined by RULE2 is regarded as the boundary of a large group. This part makes large groups from the groups made by RULE1-2. Then, when a group is made by the participants included n RULE 3, the group is one of elements in "mainstream" group or "chat" group. Figure 6 shows the case of groups. This part recognizes the group included in largest number of participants as the "mainstream" group. If some groups exist, which have the same participants, this part recognizes the group the speaker is belonged as the "mainstream" group.

2.5 Display

The display part shows the participant's name, conversation in each group on the screen. The contest of discussion is displayed in a box called "discussion box". The boxes of all groups are drawn.

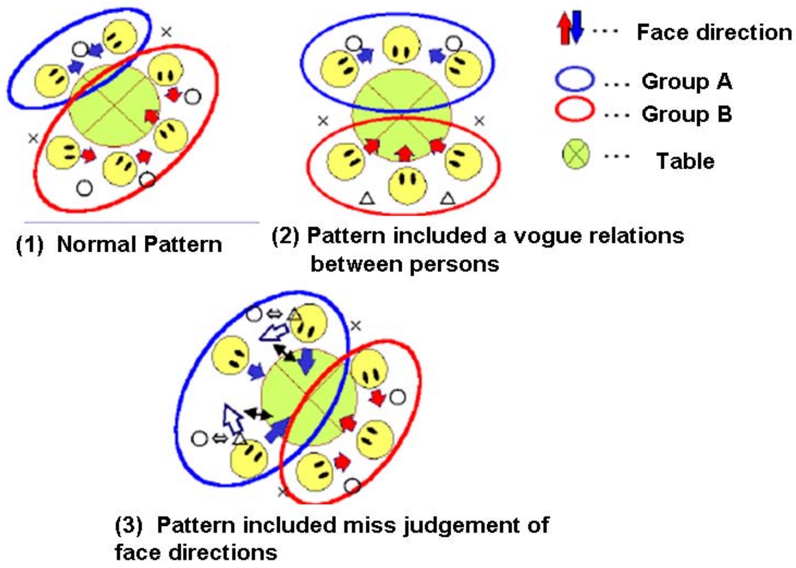


Figure 6. *Overview of some grouped situation.*

3. EXPERIMENT

This paper has an experiment to evaluate the recognition rate of face direction method.

As the experimental plan, the meeting was ready for this experiment. The participants of the meeting are 6 persons. As the situation of the meeting, 4 case studies were set up in this experiment. These case studies are shown in Figure 7.

Case (A) indicates that all participants look at the front. Case (B) shows all participants look at a next person of right side. Case (C) shows that participant of label 1 looks to the right side, participant of label 2 looks straight, participant of label 3 looks to the left side, participant of label 4 looks to the right side, participant of label 5 looks to the left side and participant of label 6 looks to the right side. Case (D) is situation that participants of label 1, 2, 3, 5 look at the front, participant of label 4 looks to the right side and participant of label 6 looks to the left side. 1,000 images are captured in each case. The face directions are extracted by using the captured images.

Image resolution, utilized by the proposed system, is 330 x 240 pixels, and the size of face areas in captured image is about 30 x 60 pixels.

3.1 Experimental Result and consideration

Figure 8 shows the face direction results of each case. The horizontal axis is the change of face direction. The vertical axis is a number of images. The proposed system recognized the face direction as the right side, when the change value is larger than 2. The system recognized the face direction as left side, when the change value is smaller than -2. The proposed system recognized the face direction as the front, when the change value is between -2 and 2.

In Case A, the graph of Case A shows that the peak of distribution of the change is zero value in all 6 persons. Therefore, the proposed system recognized the face direction as the front and the system could recognize correctly the situation of Case A in this experiment.

In Case B, the graph of Case B shows that the peaks of distribution of all subjects are larger than 2.Therefore, the result showed that the proposed system recognized that all subjects turn to right side and recognized the situation of Case B in this experiment.

In Case C and Case D, these graphs showed that the proposed system recognized the each situation.

From the results of these cases, the recognition rate of face direction was about 90 percents.

In addition, the proposed system was classified the groups by using these recognition results. As the results, in Case A, the proposed system recognized 6 subjects as one group under the recognition rate of 75 percents. In Case B, the recognition rate was 30 percents. The classification of group was low recognition rates. The face direction by using edge detection was good results to judge the direction of each person, but because the small changes in edge distributions were influenced on the classification of groups, the proposed system would be difficult to recognize the groups.

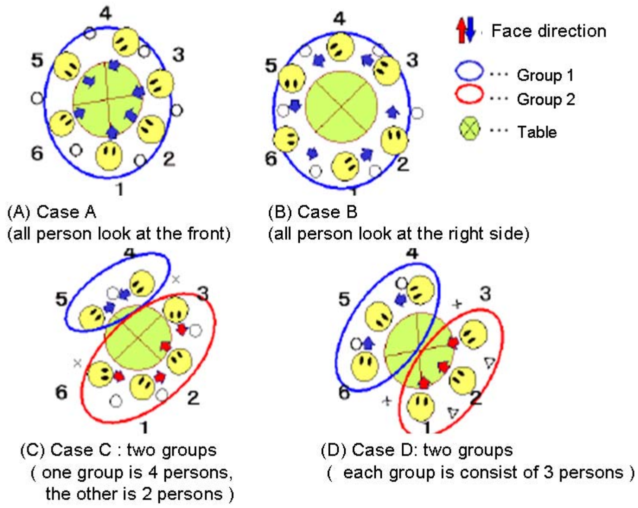


Figure 7. *Situation of Experiment.*

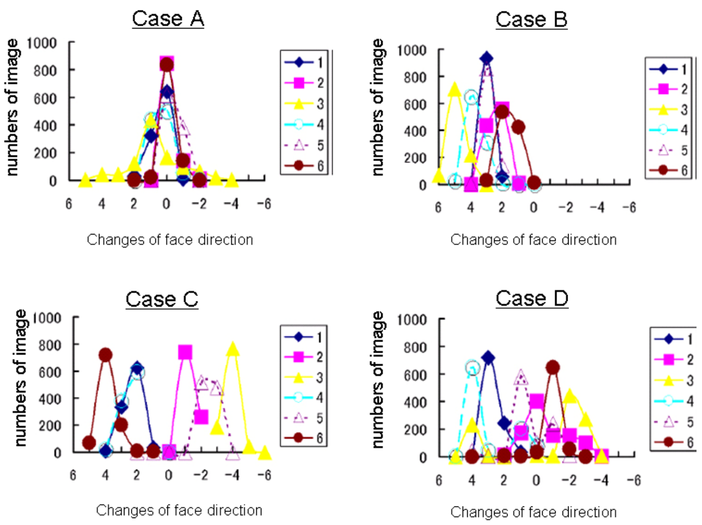


Figure 8. *Graphs of the change of face direction in each case.*

4. CONCLUSION

This paper describes a new meeting support system for the hearing impaired. The proposed system use edge distribution and skin-color-distribution to recognize a group of the mainstream of the discussion. The recognition method was evaluated in one experiment, the result showed effectiveness.

5. REFERENCES

R. Brunelli and T. Poggio(1993), Face Recognition: Features versus Templates, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15,No.10, pp.1042-1052

Y. Dai, and Y. Nakano(1995), Extraction of Facial Images from Complex Background Using Color Information and SGLD Matrices, Proc. Intl. Workshop on Automatic Face and Gesture Recognition, pp.238-242

Perceptive three dimensional interface via stereo observation

D Padbury, R J McCrindle and H Wei

School of Systems Engineering, The University of Reading
Whiteknights, Reading, Berkshire, UNITED KINGDOM

siu02dp@reading.ac.uk, r.j.mccrindle@reading.ac.uk, h.wei@reading.ac.uk

www.api.reading.ac.uk

ABSTRACT

This paper describes an intuitive approach for interacting with a computer or computer-driven applications. Interaction is achieved by observing, through a stereo camera set-up, the motion of a user's hands. This motion is then translated into 3-dimensional (3-D) coordinates to enable interaction with either a traditional 2-dimensional (2-D) desktop or a novel 3-D user interface. The aim of this work is to provide an intuitive method of interaction to computer based applications for individuals whose condition might restrict their ability to use a standard keyboard/mouse.

1. INTRODUCTION

Computer systems are now capable of producing incredible 3-dimensional (3-D) graphics but most users still work with a 2-dimensional (2-D) user interface that has remained largely unchanged for over 20 years. The orginal concept for the desktop user interface, the oNLine System (NLS), was demonstrated by Dr Doug Engelbart in 1968 (Englebart 1968). For this system a device was created for 2-D interaction which later became known as the 'Mouse'.

Today, nearly forty years on, interaction with a computer still general occurs via a combination of the keyboard and mouse. Whilst this is an appropriate form of interaction for many users, some users do not find it an intuitive form of input and for some users with mobility restrictions (e.g. lack of motor control in the fingers or arthritic conditions) they provide an obstacle rather than aid to interacting with a computer or computer-driven environment. In these instances it would be far more productive if a user could interact with the computer in an alternative manner to achieve a range of simple or complex tasks. Following on from the success of the Interaction via Motion Observation (IMO) system (Foyle & McCrindle 2004) first reported at ICDVRAT 2004, we have again used motion observation as a means of interaction but this time have extended the interaction to enable translation of the 3-D positioning of a users hands/head into a way of interacting with a 2-D desktop or a 3-D environment. To achieve this, a stereo rather than single camera set-up has been used together with a more sophisticated model of motion observation. Again like IMO, the MOTH (Motion Observation Three-dimensional Hand-tracking) system does not require the user to wear any external devices.

2. SYSTEM OVERVIEW AND APPROACH

The MOTH system uses image processing techniques to capture the 3-D position and movement of a user's hands from two cameras calibrated for stereo vision. This information is then supplied to the visualisation component of the system which is displayed as a 2 or 3-dimensional interface to the user. This process is illustrated in Figure 1.

Two Philips TouchCam II Pro cameras were used for image capture due to their high image quality, high frame rate, relatively low price and the availability of open source libraries to support development (Lavrsen 2006). C/C++ was chosen as the development language as a key requirement of the system is to respond to the user's motion in real-time. OpenCV (Intel 2006), an open source vision SDK which includes functions for camera capture, image processing and a cross platform GUI (Graphical User Interface) library was also used.

2.1 Calibration and Processing

For stereo vision to be used the system requires an initial calibration phase to derive properties of the cameras, such as their focal length and their location in the world – these are known as the intrinsic and extrinsic parameters.

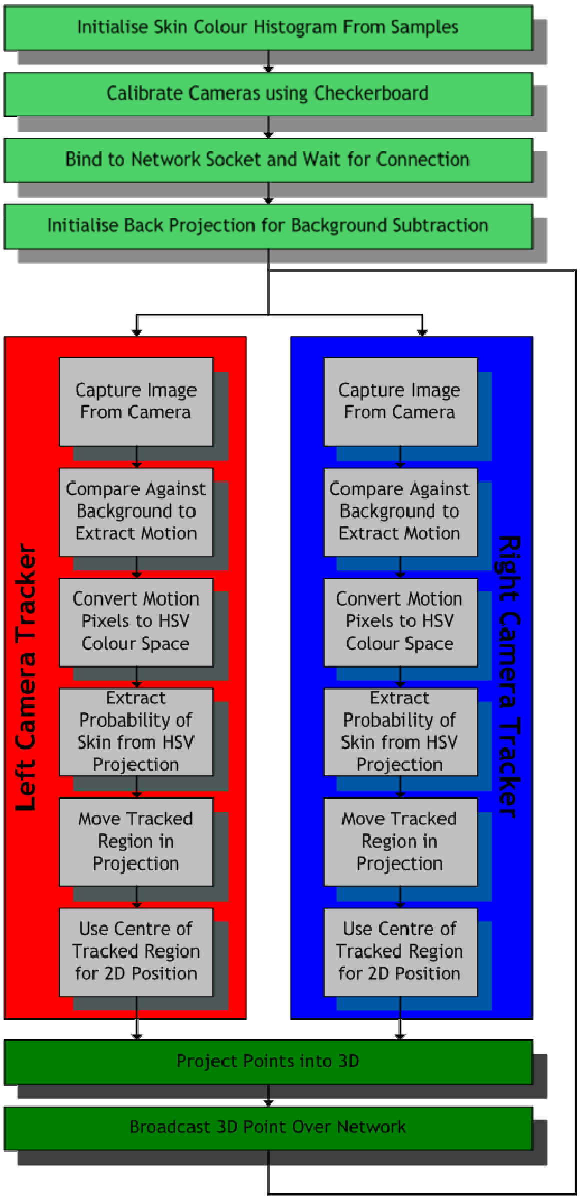


Figure 1. MOTH *System Diagram.*

Additionally, as both the computer vision tasks required for hand tracking and the three dimensional visualisation of the interface can be resource intensive, the system is spilt between two separate machines, one for the computer vision part of the system and the other for the interface. The two machines communicate over a TCP/IP channel which allows communication between different platforms to occur. In the developed system the computer vision sub-system is running on Fedora Linux whilst the interface visualisation sub-system is running on Microsoft Windows XP.

2.2 Colour Space

Computer images and graphics, and video signal transmission standards have defined many colour spaces with different properties. Some of them have been applied to the problem of skin colour modelling (Vezhnevets et al 2003). In this work, images captured from the web camera are delivered to the system in a 640 x 480 24-bit RGB (Red Green Blue) colour space but are then converted to the HSV (Hue Saturation Value) colour space. This is because whilst RGB colour values are easily corrupted by brightness, the HSV colour space separates the brightness from the actual colour (the hue). The resultant benefit of using HSV is that all people have the same skin pigment colour whatever the lightness or darkness of their skin (Bradski 1998). Brightness however can still corrupt the hue component if it is very high or very low and therefore to compensate for this all hue values

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

with a corresponding brightness within 10% of the maximum and minimum were ignored. The value of 10% was derived from our experimental experience.

2.3 Skin Colour Segmentation

Nonparametric skin distribution modelling was used in the development of the MOTH system. The key idea of this method is to estimate skin colour distribution from the training data without having to derive an explicit model of the skin colour. This method is also referred as a SPM (Skin Probability Map) (Brand & Mason 2000, Gomez 2000). By using this method, a skin colour database was established as a training data. To improve accuracy, a skin colour histogram was built using a selection of sampled skin regions from images taken from a range of users. A simple program was developed to capture an image from the web camera, and by using a basic interface to enable the user to manually select regions of skin from the captured image. Figure 2 shows an example output from the skin colour selection application.

A database of the skin colour samples was developed during the project which was then used to build a two dimensional histogram of skin colour using the hue and saturation components of the gathered skin values from a number of individuals, see Fig. 3. To achieve this, the image was first converted from RGB to HSV and then each pixel checked for a saturation outside 10% of the maximum or minimum values to avoid brightness corruption. If the pixel colour matched these criteria it was then added to the histogram of hue values. An example of image sampling and the resultant histogram is shown in Figure 3.

After training, the histogram counts are normalised to convert histogram values to discrete probability distribution as:

$$P_{skin}(c) = \frac{skin[c]}{Norm} \qquad (1)$$

where *skin[c]* represents the value of the histogram bin corresponding to the skin colour vector *c*, and *Norm* is the maximum bin value present (Zarit et al. 1999). An inequality as expressed in (2) is used as a skin detection rule (Jones & Rehg 1999).

$$P_{skin}(c) \geq \Theta \qquad (2)$$

where Θ is a threshold defined in experiments.

Figure 2. *Skin Colour Sample.*

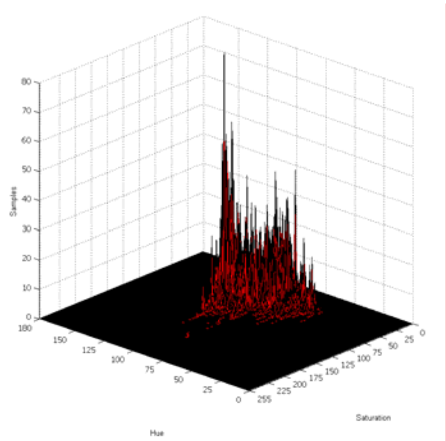


Figure 3. *Skin Histogram.*

Lighting conditions can significantly affect the results of the skin colour segmentation process. Currently all skin colour samples were captured in the same scene as the system was tested, reducing any differences in brightness. A scene with alternating lighting conditions such as windows being open/shut would make tracking significantly more difficult. An adaptive threshold technique to compensate for changes in scene brightness is currently being prototyped.

Subsequently when processing the received frame the hue component is extracted and each pixel's hue value looked up in the histogram. If it is more than a pre-determined threshold the pixel is considered to be skin coloured. This threshold is dependent on the lighting conditions of the scene and can be manually readjusted by the user during the calibration stage. The skin probability lookup process is extremely fast and produces good results, Fig. 4 shows a captured image and Fig. 5 shows which pixels of the image were considered skin coloured.

Figure 4. *Captured Image.*

Figure 5. *Detected Skin.*

Although the skin colour detection generally works well, as demonstrated above it may also recognise parts of the background as being of skin colour which would be incorrect. This false-positive is expected to be improved when adaptive thresholds are in use. In our current system skin colour combined with motion detection prevents this from occurring.

2.4 Motion Detection

The MOTH system primarily works by capturing hand or head movement and therefore a method for calculating which parts of the image have moved has been used. In this approach the first frame captured is considered to be the background and is then subtracted from all subsequent frames to calculate the movement. This approach not only captures the motion but prevents aspects of the background being mistaken for 'skin'. The motion detection process is demonstrated in Figures 6–8: the captured image is shown in Fig. 6, the current background accumulator in Fig. 7 and the detected motion in Fig. 8.

Figure 6. *Captured Image.*

Figure 7. *Background.*

Figure 8. *Detected Motion.*

2.5 Object Recognition

The techniques for skin colour recognition and motion detection are used in combination to detect the motion of an individual's hands as shown below. Fig.9 shows the image captured from the camera, Fig. 10 shows the motion detected which clearly shows the area of the hands, Fig.11 shows which pixels it believes are skin and Fig.12 shows the combination of the skin pixels and the motion. Because the users head and the skin coloured poster were both stationary only the user's hands were picked up. This method now gives a suitable back projection method for use as a tracking algorithm.

Figure 9. *Captured Image.*

Figure 10. *Detected Motion.*

 ISBN 07 049 98 65 3

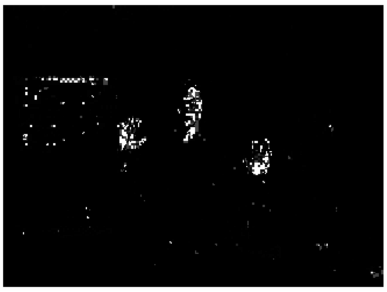


Figure 11. *Skin Colour.*

Figure 12. *Moving Skin Objects.*

The OpenCV implementation of the CAMshift algorithm with an adaptive search window size is used to find the centre of an object on the back projection (Bradski, 1984). This algorithm allows for tracking moving objects in the back projection between subsequent frames.

In this application three regions of the image are required to be tracked; the left hand, right hand and head. The position of these objects must be known before the CAMshift tracker can be started for each region. To find these initial positions three rectangles are drawn over the display from the cameras and the user is prompted to place their hand and hands in the boxes, see Fig. 13. After a short delay to allow the user to insert their hands into the regions, a CAMshift tracker is started for each region.

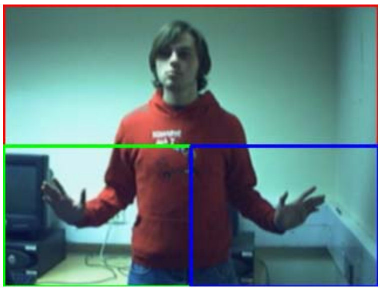


Figure 13. *Initialisation.*

The hand tracking algorithm performs very effectively, it is capable of processing 60 frames a second which provides a very smooth action even when the hands move very fast as shown in Figs. 14–17.

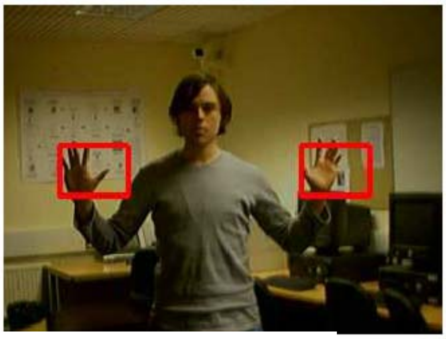


Figure 14. *Initialisation.*

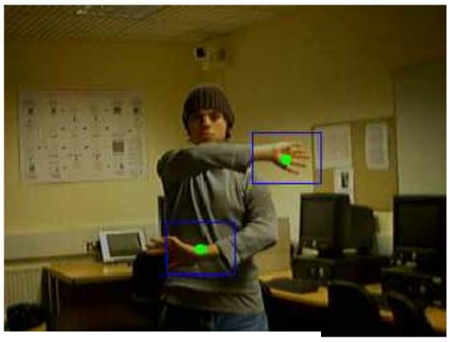


Figure 15. *Tracking 1.*

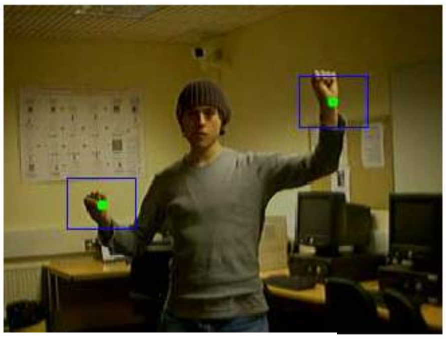


Figure 16. *Tracking 2.*

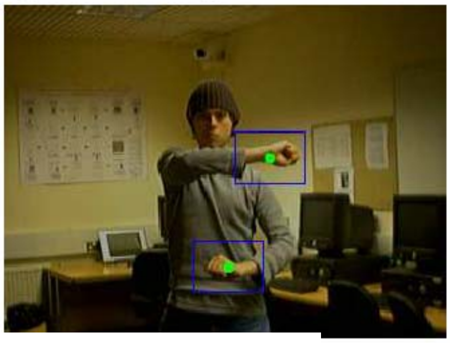


Figure 17. *Tracking 3.*

However, a potential problem which occasionally occurs, is when the object tracker loses the object it is tracking, for example, the user moves their head or hands outside of the image boundaries. A count is made of how many pixels of the tracked object are present inside the tracked region, when this decreases significantly the object is considered lost. In this even the initialisation rectangle for that object is replaced on the scene and after a short delay resumes tracking. This process is demonstrated in Fig. 18, Fig. 19, Fig. 20 and Fig. 21.

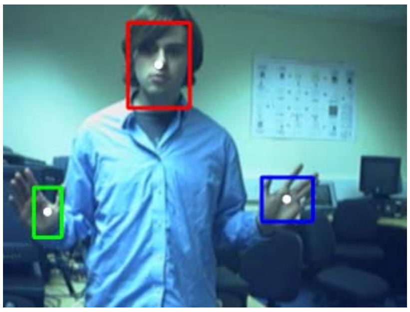


Figure 18. *Tracking Successfully.*

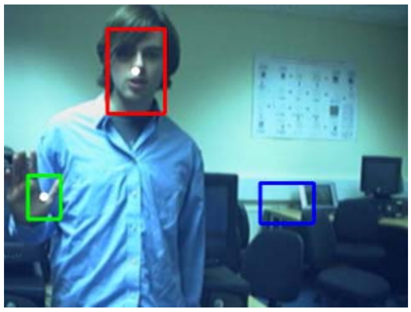


Figure 19. *Hand Removed.*

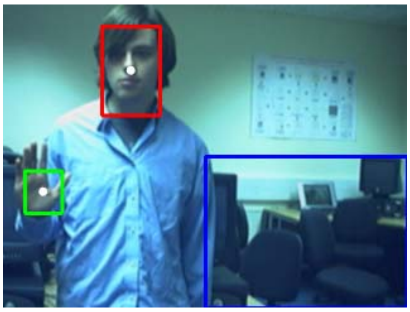


Figure 20. *Re-Initialisation.*

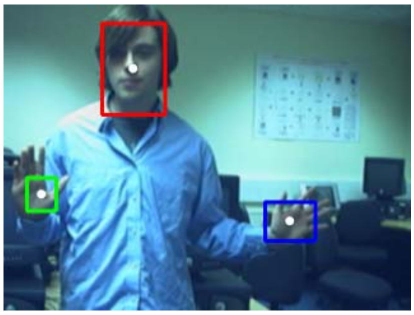


Figure 21. *Resumed Tracking.*

2.6 Stereo Vision

As stated earlier, the tracking processes are performed separately for each of the two cameras; this produces two dimensional positions for the hands and head for both cameras. The intrinsic and extrinsic parameters of the cameras are found in an initial calibration phase in which the user is required to display a checker board of known dimensions in front of both cameras. In this system the calibration procedures in OpenCV were used.

Corresponding two dimensional points in both images can be projected into a single three dimensional point in the world by finding the intersection of lines from the camera positions though the points in the image planes.

3. EXAMPLE APPLICATIONS

The above approach has been integrated into both a 2-dimensional desktop and 3-D environment. In the 2-D model the cameras can be used to detect movement of the left and right hands simultaneously. This enables more than one interaction with applications to occur at once, or to enable a simple exercise set to be developed which encourages movement of the hands in the left and right plane. An example interaction is demonstrated of the user moving both hands in a circular direction in Fig. 22; the position of the hands was tracked during motion which is shown in Fig. 23.

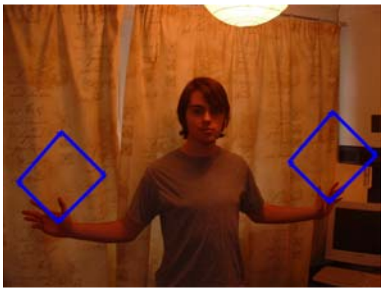


Figure 22. *Circular Hand Motion.*

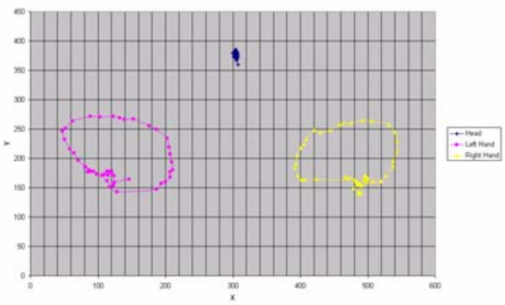


Figure 23. *Tracked Hand Motion Results.*

In the 3-D environment, movement is also encouraged in the forward and back plane, for example by manipulating a number of coloured spheres within the 3-D space, see Fig. 24. This can become a more entertaining exercise by placing a 3-D tube in the environment, such that the spheres can be controlled though the tube enabling simple basketball style games to be played. Fig. 24 shows the user manipulating the 3-D spheres.

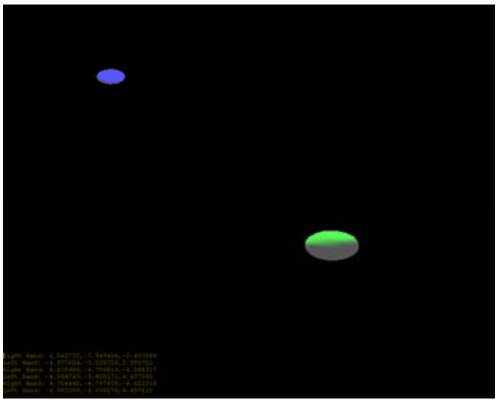


Figure 24. *3-D Ball Exercise.*

Figure 20. *System Use.*

4. FURTHER APPLICATIONS

There are a number of areas to which this technology and the MOTH system can be applied. These include the ability to interact with a traditional 2-dimensional computer desktop in a more productive manner and without the need to type or control the mouse; the ability to control a range of devices such as MP3 players, televisions or radios; and the ability to interact with objects in a 3-dimensional interface in order to develop spatial awareness or to coach or encourage a user to perform certain exercises for general health or rehabilitation purposes. A number of users have had the opportunity to use the MOTH system and results to date have been encouraging. The next stage is to conduct more formal trials with users who have limited fine motor control and to develop more fully the 3-dimensional aspects of the interface and corresponding exercises and games.

5. CONCLUSIONS

This paper has described a new way in which individuals, who do not possess the fine co-ordination required to use a mouse or keyboard, can interact with a computer or computer-driven environment. The MOTH system has been shown to effectively detect hand and/or head movements and translate these movements into inputs to a 2-D or 3-D interface. The interface may be populated with traditional computer desktop applications or may run applications which through a pattern of motion would encourage a user to undertake general or rehabilitation-oriented exercises, or to enable them to control their environment.

6. REFERENCES

G R Bradski (1998), *Computer Vision Face Tracking For Use in a Perceptual User Interface.*

J Brand and J Mason (2000), *A comparative Assessment of Three Approaches to Pixellevel Human Skin-detection*, Proc. of the 16th International Conference on Pattern Recognition, vol. 1, pp 1056-1059

D Engelbart (1968), *Demo of the Graphical User Interface*, http://sloan.stanford.edu/mousesite/1968Demo.html

M Foyle and R J McCrindle (2004), *Interaction Via Motion Observation (IMO)*. ICDVRAT 2004.

G. Gomez, 2000, *On Selecting Colour Components for Skin Detection*, Proc. of the 16th International Conference on Pattern Recognition, vol. 2, pp 961-964

Intel (2006) *Open Source Computer Vision Library*, www.intel.com/technology/computing/opencv/index.htm.

M J Jones and J M Rehg (1999), *Statistical Colour Models with Application to Skin Detection*, Proc. of CVPR'99, Vol. 1, pp. 274 – 280.

K Lavrsen (2006), *PWC Documentation Project*, http://www.lavrsen.dk/twiki/bin/view/PWC/WebHome.

V Vezhnevets, V Sazonov and A Andreeva (2003), *A Survey on Pixel-Based Skin Color Detection Techniques,* Proc. Graphicon-2003, pp. 85-92, Moscow, Russia, September 2003.

B D Zarit, B J Super and F K H Quek (1999), *Comparison of Five Colour Models in Skin Pixel Classification*, Proc. of ICCV'99 International Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-time Systems, pp58-63.

Evaluation of a computer aided 3D lip sync instructional model using virtual reality objects

A Rathinavelu[1], H Thiagarajan[2] and S R Savithri[3]

[1]Dr Mahalingam College of Engineering and Technology,
Pollachi-642003, TamilNadu, India,

[2]National Institute of Technology,
Trichy, TamilNadu, India

[3]All India Institute of Speech and Hearing,
Mysore, Karnataka, India.

[1]*starvee@yahoo.com*

ABSTRACT

Lip sync model is one of the aspects in computer facial animation. To create realistic lip sync model, a facial animation system needs extremely smooth lip motion with the deformation of the lips synchronized with the audio portion of speech. A deformable, parametric model of the lips was developed to achieve the desired effect. In order to create realistic speech animation, the articulatory modeling of the lip alone is insufficient. The other major articulators such as the tongue and jaw must also be considered. This lip sync model was initially developed by using polygonal model and blended key shape techniques and then parameterized by using 36 control points. The data for lip sync model was collected from video image and magnetic resonance imaging (MRI) techniques. The articulatory movements of our lip sync model were presented along with virtual reality (VR) objects in an interactive multimedia (IMM) interface. This IMM interface was used to teach small vocabulary of hearing impaired (HI) children. Virtual reality objects used to increase the cognitive process within HI children. The bilabial speech sounds were differentiated by using appropriate visual cues. Control panel was developed to present articulatory movements at different speed. For this study, six hearing impaired children were selected between the ages 4 and 7 and they were trained for 10 hours across 2 weeks on 18 meaningful words. The intelligibility of hearing impaired children was experimented to find out their performance in articulation and in memory retention. The results indicated that 65-75% of given words were articulated well and 75-85% of words were identified by all children.

1. INTRODUCTION

Once a child has a reasonable command of language and his phonetic and phonologic skills enable him to produce most speech patterns (including some consonant blends), it becomes possible to measure the intelligibility of his spontaneous speech [Ling, 1976]. An articulation disorder of hearing impaired children may be defined as incorrect production of speech sounds due to faculty placement, timing, direction, pressure, speed or integration of the movements of lips, tongue, velum or pharynx. A teacher/clinician should be able to evaluate the child's performance in articulation of speech sounds by employing suitable computer aided speech training system. It is assumed that the child's vocabulary and his knowledge of the lexical, morphological, syntactic and semantic rules which are essential to meaningful speech communication [Ling, 1976]. To keep track of the variety of individual children's specific speech training needs, teacher must spend lot of time to prepare the log file in the form of chart (or) progress card. In order to minimize the time involvement of teacher, we suggest computer aided sub skills for speech training needs of hearing impaired (HI) children. The speech of hearing impaired children differs from speech of Normal Hearing (NH) children in all aspects [Ling, 1989]. According to deaf researchers [Lundy, 2002], hearing children have consistently demonstrated the ability to perform such tasks between the ages of 4 and 5 years. But hearing impaired children are delayed by approx. 3 years in this cognitive developed milestone. It is also found that hearing impaired children are significantly delayed in the development of language skills. It has been known for some time that hearing impaired children can make use of speech reading for acquisition of words. To speech read, the children must observe the teachers articulatory movements of the lips, jaw and the tongue. But many children fail to observe the articulatory movements of a teacher [Rathinavelu, 2003].

Due to auditory degradation of HI children, perceiving visual info is most useful informative for them. Computer Aided Articulation Tutor (CAAT) may be an alternative one for the hearing impaired children to acquire speech sounds and syllables by perceiving visual info without teacher's assistance [Rathinavelu, 2003]. Incorporating text and visual images of the vocabulary to be learned along with the actual definitions and spoken words, facilities learning and improves memory for the target vocabulary [Massaro & Light, 2004A]. According to empirical findings, children are good at producing spoken language if they do better at speech perception. When language and articulation disorders coexist, treating only one of them may produce some effect on the other, but it is likely that the effects will not be substantial; more research is needed [Pena-Brooks,2000]. Children who misarticulate may have additional problems in overall language skills. Our approach is a fairly new concept of using computer aided lip sync model using 3D VR objects to present articulatory position of complex speech sounds of Tamil language.

The lip sync model visualizes the articulator movements as described by the articulation parameters like lips, jaw and tongue. If gestures and speech express the same meaning, then gestures and speech should function as two sources of info to be integrated by the perceiver [Massaro, 1998]. Both visual and auditory cues are very important for children to acquire speech sounds. Previous studies [Massaro, 1998] indicated that visual stimuli are integrated into perception of speech. In Tamil, there is no exclusive study conducted in the construction of an articulatory model and there is no development of 3D articulatory-animated model for bilabial speech sounds and syllables. Few multimedia researchers [Massaro, 1998; Rathinavelu, 2001 and 2003; Barker, 2003] suggested about presenting new words graphically for hearing impaired children. So they are able to acquire much faster than conventional book form.

2. DESIGN OF LIP SYNC ARTICULATORY MODEL

While the auditory signal alone is adequate for communication, visual information from movements of the lips, tongue and jaws enhance intelligibility of the acoustic stimulus. Adding visible speech can often double the number of recognized words from a degraded auditory message [Massaro & Light, 2004B]. Visual pattern is particularly more effective in recognition of articulatory position of speech sounds for Hearing Impaired (HI) children. Articulatory process helps to convert the linguistic messages into sound. The most fundamental research issue of articulatory process is how to train the speech segments of a language. Polygonal model was chosen here to develop 3D lip sync model and animation so naturalistic.

Lips and tongue movements are the most powerful objects with several skeletons created for the realistic animation. Video clips of speakers articulating words were shot and the frames of this footage provided a guide for the corresponding frames of an animation [Lewis, 1991]. The lip sync model was constructed by digitizing the video image using Polygonal modeling. A 2D surface based coordinate grid was mapped onto the front and side images of a face. Point correspondences were established between the two images and the grid was reconstructed in 3D space [Parent, 2002]. 3D Animation involves three steps: Modeling, animation and rendering. An articulated model is a collection of many deformable objects connected together. The 3D lip sync model has been created with correspond to the visemes that constitute a word (refer figure.1).

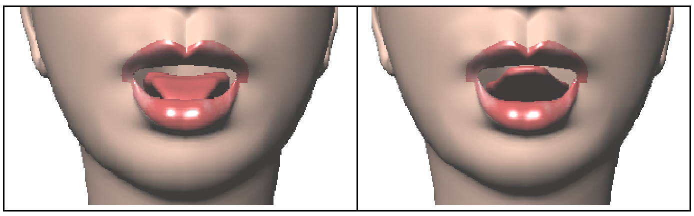


Figure 1. *Lip Sync model.*

These models serve as key frames for subsequent animation. Key framing allows fine control to ensure naturalistic facial movements of lip sync model. Interpolation between a finite set of visual targets was used to achieve speech articulation [Bailly, 2003]. Natural speech of speaker was then synchronized as realistic audio visual sequence. The articulation of voiced speech sounds were indicated by coloring the vocal cord portion of model. Table 1 shows about the parameters used to develop our lip sync model.

Table 1. *Parameters of lip sync model*

No	Articulatory Features	Remarks
1	Voiced or unvoiced	Graphically Highlighted
2	Opening and Closing of lips	Video image Data
3	Rounding of the lips	Video image Data
4	Tongue Position	MRI Data
5	Jaw Movement	Video + MRI Data

The tongue movements were built by animating tongue raise, tongue contact (with palate) and tongue curved. The data for tongue modeling was retrieved from MRI of natural talker [Rathinavelu]. The corresponding phonemes were matched with movements of mouth. In general, articulation of each given word is built from moving articulator's lips, jaw and tongue. In our model, lips and tongue articulators (including jaw) were integrated together to work as a single speech production model in an IMM interface (refer figure 3). The lip was viewed as being made up of 24 polygons each consisting of 5 vertices. The 36 control points, 12 from each of the outer, middle and inner lip contours constitute the vertices of the polygon. The specified polygons were triangulated and then rendered with either Material or Wire frame appearance. Suitable lighting and shading effects provided a realistic appearance of the lip. Our parameterized lip model was obtained by exporting the data from the Image based 3D model. Data collection, labeling and storing was a challenging task to develop this articulatory model (refer Table 2).

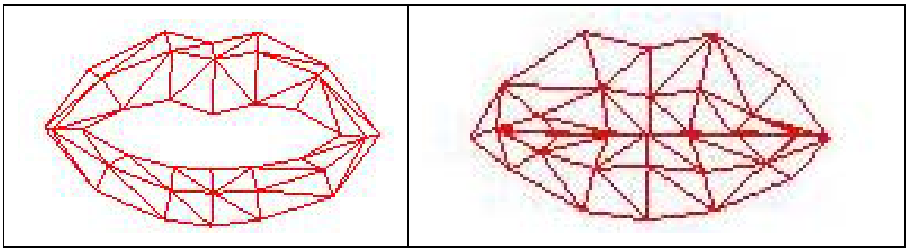


Figure 2. *Parameterized lip model for 'pa' and 'm'.*

The desired animation was obtained by specifying the target control points through a control panel interface. Using this parameterized model, the articulatory movement of the lips was modeled for bilabial Tamil speech sounds such as 'pa', 'mm' etc. as shown in figure 2.

Table 2: *Sample of data for speech sound 'pa' in the word 'palam (fruit)'.*

Point Number	**X Coordinate**	**Y Coordinate**	**Z Coordinate**
1	2.58766	–0.015768	0.04
2	1.93075	0.85171	1.11022e–016
3	1.08925	1.27246	1.11022e–016
4	0.5002	1.13782	1.11022e–016
5	–0.08885	1.28929	1.11022e–016
6	–1.04816	0.80122	1.11022e–016
7	–1.62247	–0.236976	–0.19
8	–0.962	–0.7148	0
9	–0.095	–1.091	0
10	1.095	–1.091	0
11	2.013	–0.794	0
12	1.80066	–0.435607	0.65

In order to create realistic speech animation, the articulatory modeling of the lip alone is insufficient. The other major articulators such as the tongue and jaws are also suitably modeled. The figure number.3 represents the articulation of the Tamil word 'palam'. The phonemes 'pa' and 'm' were modeled using the lips alone. The phoneme 'la' involved the tongue; hence the articulation of 'la' was modeled from MRI data of Speaker [Rathinavelu].

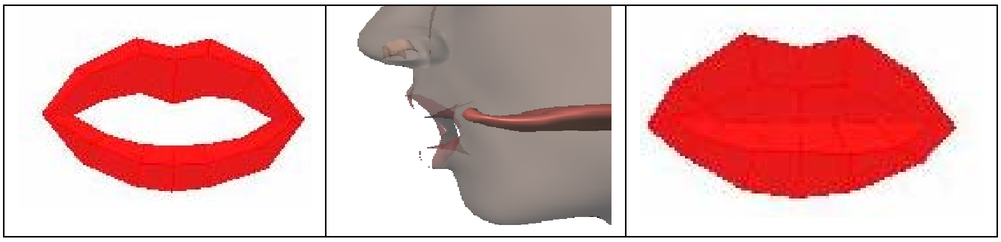


Figure 3. *Parameterized lip model along with tongue portion for the word 'pa-la-m'.*

Virtual reality object gives 360 degree view and more cognitive processes within children to retain newly learned activities [Rathinavelu, 2006]. The making of VR objects can be divided into the following phases: drafting, modeling, animation and exporting into VR Scene. Drafting software was used to draw an object from dimensions of the video image of real-world object and then polygonal modeler was used to model 3D view of the object. VR software helped to view the object three dimensionally after animating them suitably.

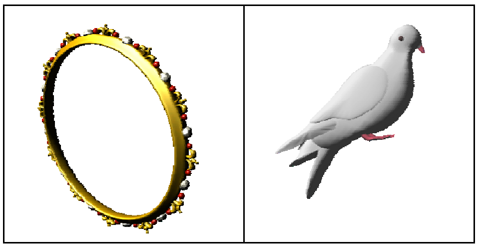


Figure 4. *Three dimensional Virtual Reality Objects.*

The interface of CAAT displays both the articulatory position of each word and 3D view of corresponding object together. Interactive Multimedia (IMM) software helped to reduce stress and strain of teachers who involved in content preparation and replicated teaching for specific speech sounds/syllables [Rathinavelu, 2006].

3. METHOD

The goal of our study was to examine the effectiveness of a Computer Aided Articulation Tutor (CAAT) to teach the articulatory position of speech segments for hearing impaired Children. Prior to scheduled training, teachers were involved to select the unknown but meaningful words under the categories of bilabial speech sound p, b, and m. The combination of consonants [p, b, and m] and vowels [a, e, I, o, and u] helped us to build 18 meaningful words for this computer aided articulation test. As shown in table 3, three sets of 6 words each were created. There were three stages in this study: an initial test, training for 2 weeks and a final test.

The training interface of CAAT was designed to combine lip sync model and Virtual Reality Objects to engage hearing impaired children effectively. Consonants have been traditionally described according to place, manner, and voicing dimensions. Place of articulation indicates where along the vocal tract the consonant is formed; manner of articulation indicates how it is formed; and voice indicates whether the vocal folds are vibrating during its production [Pena-brooks, 2000]. To avoid the confusion in perceiving the bilabial sounds 'p' and 'b', visual cues were used in IMM interface to differentiate 'p' and 'b'. For example, the visual cue 'palam' (fruit) was used for speech sound 'p' and the visual cue 'ball' was used for teaching syllables starting with 'b'. The Lip sync model was built-in with four modules as described below.

- Basic research module: A new concept was initiated to design and develop computer aided lip sync model using VR objects to help HI children in the acquisition of articulatory movements of bilabial speech sounds [p, b, and m] and syllables.
- Development module: Data of video image and MRI was used to develop 3D lip sync model. The lips were modeled using 36 control points distributed equally over three contours – outer, middle and inner. This module involves data collection, labeling and storing.
- Training module: 18 meaningful words were presented to 6 hearing impaired children for 10 hours over 2 weeks. Three boys and three girls were selected as subjects. Every day, they spent 30 min each in

beforenoon and afternoon sessions. Six words were selected in each category of bilabial speech sounds p, b, m.

- Evaluation module: After 2 weeks training, Children were asked to articulate the speech sounds and syllables on 10th day. The results were aimed to discuss about their performance in articulation and memory retention. In IMM interface, VR objects only were presented to identify the suitable/correct meaningful word from the list of words displayed.

[p] In its production the lips are closed and the soft palate is raised to close the nasal passage. When the lips are opened the air suddenly comes out with explosion. There is no vibration in the vocal cords. This sound may be described as a voiceless bilabial stop [Rajaram, 2000].

[b]. The movements of the speech organs are exactly the same as those for its corresponding voiceless variety[p] except for the vibration of the vocal cords. In Tamil it occurs initially in some of the borrowed words and medially after the nasal [m]. This may be described as a voiced bilabial stop [Rajaram, 2000].

[m]. In its production the lips are closed. The soft palate is lowered and the air stream comes freely through the nasal cavity. The vocal cords are vibrated during its production. In short, the articulatory movements for [m] are the same as for [b] except that the soft palate is lowered and the air is emitted through the nasal cavity. So the formation of this sound may be described shortly by defining it as a voiced bilabial nasal [Rajaram, 2000].

Table 3. *Meaningful words chosen for test.*

p	b	m
Pu-li (tiger)	Bim-bam (image)	Ma-ra-m(tree)
Pu-l(grass)	Kam-bam (post)	Mak-kal(people)
Pa-la-m (fruit)	Bak-ti (devotion)	Miin (fish)
Up-pu (salt)	Am-bu (arrow)	buu-mi (earth)
Kap-pal(ship)	Cem-bu (copper)	Paam-bu(snake)
Ap-paa(father)	Ba-la-m (strength)	Am-maa(mother)

4. EXPERIMENTS AND RESULTS

Our evaluation aimed to determine the degree to which the CAAT contributes to the acquisition and retention of meaningful words as well as student's response to our lip sync model. According to training schedule (refer table 4), 18 words were gradually presented to them during 2 weeks period. At that time, participants were asked to perceive the articulatory movements of Lip Sync Model to identify the speech segments of each word presented to them. Teachers and parents were instructed not to teach or use these words. During the training period of 2 weeks; it was still possible that the words could be learned outside of this training environment.

Table 4. *Training schedule.*

Day	**Weekly Schedule**	
	First week— Set of words	**Second week — Set of words**
1	1	3, 1
2	2	1, 2, 3
3	3	2, 3, 1
4	1, 2	3, 1, 2
5	2, 3	Test

During two weeks period of training, every child was trained by until they achieve maximum accuracy in articulation of each word and then only they were instructed to move to next word. No feedback was given to them during training period. Testing and training were carried out in a quiet room at oral deaf school, for about 60 min each day of a week for two weeks. After training period, students were instructed to articulate the words on 10th day. At that time, the lip sync face was removed from the interface of CAAT. Instead, 3D VR object and few similar titles of object were presented for each word. The children were instructed to recognize, identify and articulate the words one by one in front of CAAT without any guidance from teachers. The performance of each participant was rated as good and poor (Ref. figure 5).

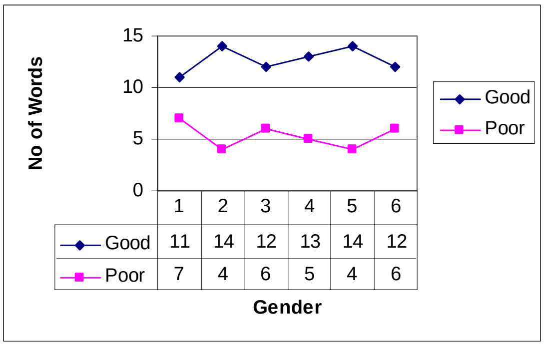


Figure 5. *Performance of student's articulation.*

Even after two weeks training, few words were not articulated properly by some of the hearing impaired children in the final test. For example, 'uppu' (salt) was substituted with 'o' and articulated as 'oppu'. This was due to similar kind of lip protrusion perceived by most of the children. Few of them didn't have any control over short vowels. For example, 'pu-li' (Tiger) was articulated as 'puu-li' and 'pu-l' (grass) was articulated as 'puu-l'. The Common articulatory errors were noted as detailed below.

Table 5. *Substituted speech sounds by few HI children.*

Substituted Speech sounds	Target Speech Sounds	Articulated Speech Sounds
/bi/-/pi/	Bi-m-bam	Pi-m-bam
/pa/-/ba/	Pa-lam	Ba-lam
/ba/-/pa/	Ba-k-ti	Pa-k-ti
/ra/-/la/	Ma-ra-m	Ma-la-m

Apart from these substitution errors, most of the children faced difficulty to articulate the words like 'ma-k-kal' (people) and 'buu-mi' (earth). It was commented by teachers that those words are meaningful but not familiar for hearing impaired children to use. This kind of incorrect articulation was considered as poor performance. This investigation also helped us in obtaining information about articulatory-acoustic performance of bilabial speech sounds and syllables.

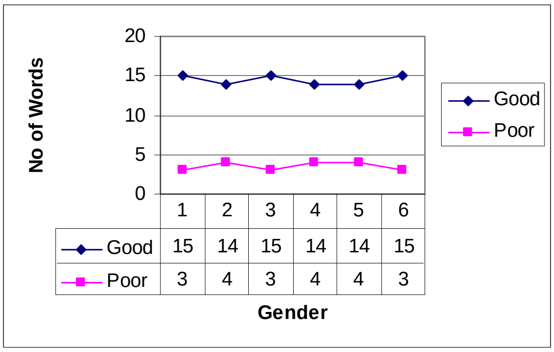


Figure 6. *Performance of student's memory retention.*

In memory retention test, all the children did well in identifying the VR object with suitable title of the particular VR object in the given list. As indicated earlier, some of the unfamiliar words were not retained well by most of the children. Those words were like 'mak-kal' (people), 'bim-bam' (image), 'pus-bam' (flower), 'Buu-mi' (earth). Rest of the words was identified by almost all the children. The individual articulatory performance of hearing impaired children is shown in figure 5 and their performance in memory retention of words is presented in figure. 6.

5. CONCLUSION

Our goal was to develop computer aided articulatory model to help Hearing impaired children in acquisition of speech sounds and syllable. Instead of teacher's role, computer can do tireless job by repeating the articulatory movements any number of times. It has been proved in our computer aided lip sync model that the visual information is most useful for the children with hearing loss in acquisition of speech sounds and syllables. Previous research findings suggested [Massaro, 1998] that hearing impaired children were able to response well when the appropriate visual cues (text, image and graphics) are presented to them. Our lip sync model integrates lips and tongue movements to articulate the given words. This model helped hearing impaired children to perceive the position and manner of each speech sounds of the word. Since co-articulation is an important issue with perception of speech segments with hearing impaired children, the control panel was used to show articulatory movements at different speed. Substitution and addition disorders were noted as most common errors during articulation of hearing impaired children. It was also noted that the software interface had really motivated and helped most of the children to correct their articulatory errors by them. Every child was given two weeks training and then tested with our lip sync model to identify their performance in articulation and memory retention. The results indicated that 65-75% of given words were articulated well and 75-85% of words were identified well by all children.

Acknowledgements: This research work is financially supported by Dr Mahalingam College of Engineering and Technology, Pollachi, South India. I wish to thank members of college management and faculty collegues of CSE and IT for their extended support.

6. REFERENCES

G Bailly et al (2003) , Audio Visual Speech Synthesis, International Journal of Speech Technology, Kluwer Academic Publishers, Netherlands, pp.331–346.

L J Barker (2003), Computer-Assisted Vocabulary Acquisition: The CSLU Vocabulary Tutor in Oral-Deaf education, Journal of Deaf Studies and Deaf Education, 8(2), spring 2003, pp.187–198.

J Lewis (1991), Automated Lip-Sync: Background and Techniques, The Journal of Visualization and Computer Animation, 2:118–122.

D Ling (1976), Speech and the hearing impaired child: Theory and Practice, Alexander Graham Bell Association for the deaf, Inc, U.S.A.

D Ling (1989), Foundations of spoken language for hearing impaired child, Alexander Graham Bell Association for the deaf, Inc, U.S.A.

J E B Lundy (2002), Age of language skills of deaf children in relation to theory mind development, Journal of deaf studies and deaf education 7:1, PP.41–56

D W Massaro (1998), Perceiving Talking Faces: From speech Perception to a Behavioral Principle, MIT Press, U.S.A

D W Massaro & J Light (2004A), Using visible speech for training perception and production of speech for hard of hearing individuals, Journal of Speech, Language, and Hearing Research, 47(2), pp.304–320.

D W Massaro & J Light (2004B), Improving the Vocabulary of children with hearing loss. Volta Review, 104(3), pp 141–174.

R Parent (2002), Computer Animation Algorithms and Techniques, Morgan Kaufmann Publishers, San Francisco.

A Pena-Brooks and M N Hegde (2000), Assessment and treatment of articulation and phonological disorders in children, Pro-Ed Publishers.

S Rajaram (2000) Tamil Phonetic Reader, CIIL, Mysore, India.

A Rathinavelu (2003), Early language learning for elementary students with hearing impairment in India, Masters' Thesis, School of Communications and Multimedia, Edith Cowan University, Australia.

A Rathinavelu, A Gowrishankar (2001), e-Learning for hearing impaired (p21.1–21.6), Proceedings of the Apple University Consortium Conference 2001 (ISBN 0-947209-33-6), James Cook University Townsville, Queensland Australia, Sept 23–26.

A Rathinavelu et al ,Interactive multimedia tool to help vocabulary learning of hearing impaired children by using 3D VR objects as visual cues, National Journal of Technology, (Accepted for publication).

Constructing new coordinate system suitable for sign animation

T Kuroda[1], K Okamoto[2], T Takemura[3], K Nagase[4] and H Yoshihara[5]

[1,3,4,5]Department of Medical Informatics, Kyoto University Hospital,
54, Kawahara-cho, Shogo-in, Sakyo-ku, Kyoto, JAPAN

[1,2,3,4,5]Graduate School of Informatics, Kyoto University,
36-1, Yoshida-honmachi, Sakyo-ku, Kyoto, JAPAN

[1,2,3,4,5]*{Tomohiro.Kuroda, kazuya, takemura, nagase, lob}@kuhp.kyoto-u.ac.jp*

www.kuhp.kyoto-u.ac.jp/~mi/

ABSTRACT

This paper proposes new coordinate system suitable for denoting sign language motion. As the proposed coordinate system consists of polar coordinate systems whose origins are certain points of human body, postures shown on the system can be proportional for avatars with any possible shape and fit with existing subjective sign notation systems. This paper extracted coordinate origins from Japanese-Japanese Sign Language Dictionary via morphological analysis. Selected 85 points are successfully mapped on H-ANIM standard humanoid avatar.

1. INTRODUCTION

Sign linguistic engineering is a group of research to develop communication aid (called sign information system) for the Deaf and the Hearing Impaired, who have communication barrier in social lives, using information technology (Nagashima and Kanda 2001). Figure 1 shows the general view of sign information systems. Any sign information systems, such as virtual reality based sign language telephones, automatic translation system between sign languages and phonetic languages, and even sign language dictionary, share same basic structure. Most of sign information systems use three-dimensional computer graphic model of human beings called avatar to show signs. To obtain signs, though several systems using image recognition technologies, most of them are utilizing motion capture systems consists of data-gloves and position and orientation sensors either magnetic-field-based, ultrasound-based, and image-based. The authors also developing sign language telephone named S-TEL (Kuroda et al 1998) and consequent systems based on motion capture system consists of two data-gloves and three position and orientation sensors.

Most of sign animations handles motion as sets of rotation angles of each joint and generate animation applying the data to an avatar model, just as same as MPEG-4 H-ANIM standard (ISO/IEC, 2001). However, a certain body motion parameter to show a certain sign on a certain avatar cannot produce the same sign when it applied on another avatar with different body shape. Thus, the sign linguistic system needs to prepare an avatar, which holds the same body shape as a certain person to produce proper signs. Although authors developed a system to morph a standard avatar to fit a certain person through image recognition (Kuroda et al, 2001), it cannot give users full selection of avatar's outlook to have favorable view. Thus, to give flexibility and usability on sign information systems, a new motion coding system, which is independent from body shape and easily transferable from or to standard motion data such as a captured motion, data via a certain motion capture system or H-ANIM.

This paper proposes new coordinate system suitable for sign animation. Notations of textbooks or dictionaries of sign language present a position and orientation of hands in a certain posture in relation to a certain point of human body.

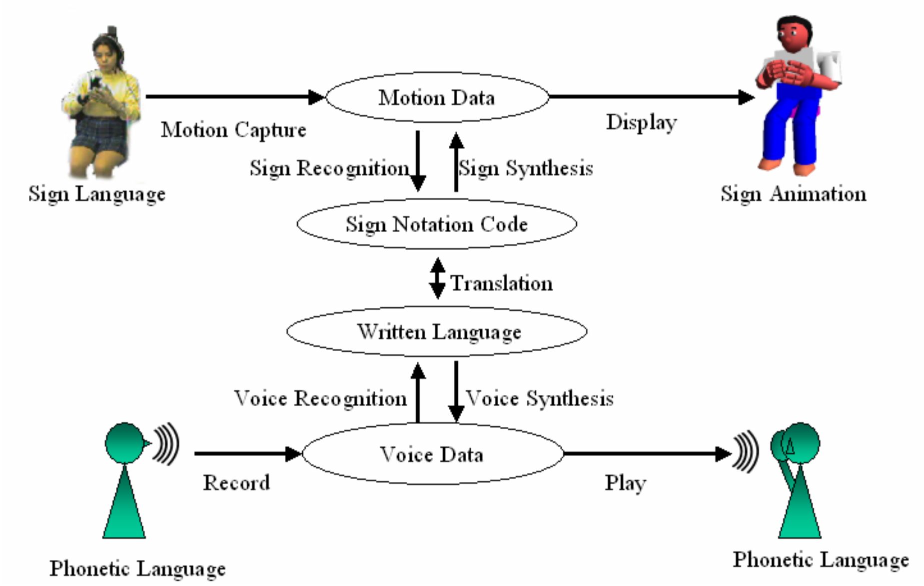


Figure 1. *General overview of sign information systems.*

2. FOREGOING CODING SYSTEMS

Foregoing researches on sign language developed several types of sign coding systems. Most of the sign notation system follows framework of Stokoe (1965), which claims signs are consists of "Cheremes", that is, hand posture (Dez), position (Tab) and motion (Sig). Thus, most of the notation system denotes signs as a series of combination of hand posture, orientation and position, although some of them denote motions as shape of trajectories. Sometimes several non-manual signals, such as facial expressions, are added on them.

HamNoSys (Hanke, 2004), one of the most well-known sign notation system, also denote signs as a combination of hand posture, orientation and position. HamNoSys denote position of hands as combination of upper/middle/bottom of the shoulder and left and right of body, although it adds several more codes on update of version 4.0, the resolution is still quite rough.

SignWriting (Satton, 1998), another one of the most well-known sign notation system, denotes hand posture and position, orientation and speed. As SignWriting is designed for easy denoting and reading of sign language and is derived from dance notation system, the notation is quite pictorial and subjective. Thus, automatic conversion from or to motion capture data seems quite difficult.

Kurokawa (1992) developed a motion coding system for nonverbal interface and applied the coding system to denote sign language. However, resolution of the coding system, again, quite rough to synthesis signs animation from given code as the system designed mainly for motion recognition.

As discussed above, foregoing sign notation systems are based on subjective coding and its resolution are not detailed enough, new coding system, which mediates such subjective coding system and motion data, such as FACS (Ekman and Freisen, 1978) for facial expression, is essential for sign language information systems. As hand posture can be mapped via simple vector quantization as authors presented on previous conference and implemented on StrinGlove® (Kuroda, 2004), the coding system to denote hand position and orientation can be sufficient.

3. DESIGNING NEW COORDINATE SYSTEM

3.1 Basic approach

As discussed above, the target of this research is to develop hand posture/orientation coding system to mediate raw captured or kinematics data and existing subjective sign notation codes.

Careful observation of foregoing sign notation systems, especially new location symbols added on HamNoSys in version 4.0 (Schmaling and Hanke, 2004), gives interesting insight of sign posture coordination. Figure 2 shows several new added location symbols. As shown in the figure, new version of

Proc. 6th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Esbjerg, Denmark, 2006
©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

HamNoSys gives locations of nominal hand by several body parts such as lip, ear lobe, nose and arms. On the other hand, quick review over several textbooks of sign languages also gives positions of hands or fingers by referring certain body parts. Therefore, human being seems to position hands relative to a certain body parts, such as lip, nose, chest, elbow, etc. Thus, coding system that let us denote relation between hands and other body parts can fit subjective insight of human beings as well as existing sign notation code.

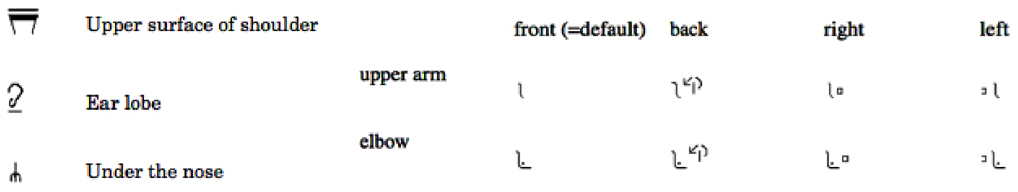


Figure 2. *New location symbols on HamNoSys 4.0 and its usage.*

On the other hand, to enable automatic conversion from or to “raw” motion data, such as captured motion data or generated kinematics data, the coding system must not be sparse or rough. The coding system should have enough resolution to show its precise position in a certain coordinate.

The authors propose to develop new coordinate system suitable for sign motion rather than coding system. The coordinate system consists of sets of polar coordinate system whose origins are certain points (the reference points) on human body, such as “inner corner of right eyelid” or “eye”. As the reference points are proportional to shape of body, the proposed system can proportion position and orientation of hands to various avatar models including cartoon-like model with big head as shown in Fig. 3.

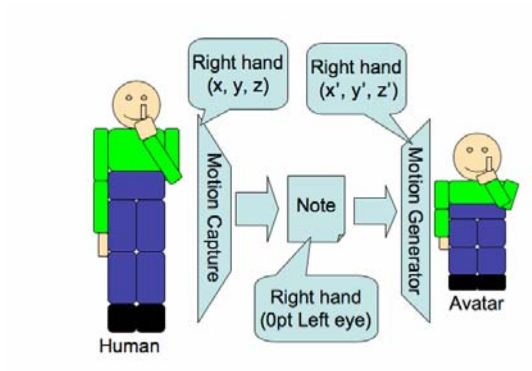


Figure 3. *Basic concept of the proposed coordinate system.*

This coordinate system allows automatic conversion from “raw” motion data by calculating out distance of hand from several reference points and to “raw” motion data by mapping them on certain avatar. Additionally, a certain motion denoted on subjective notation code can be mapped onto the coordinate by giving distance on a certain rule.

3.2 Picking up the reference points

To obtain reference points, this paper analyzes notations of Japanese – Japanese sign language dictionary issued by the Deaf association of Japan (Yonekawa, 1997). The dictionary has 8320 headwords (Japanese) and one headword has a pair of gestures (Japanese signs) and notation of each sign gestures in Japanese as shown in Fig. 4. In total, the dictionary has 16640 sign notations.

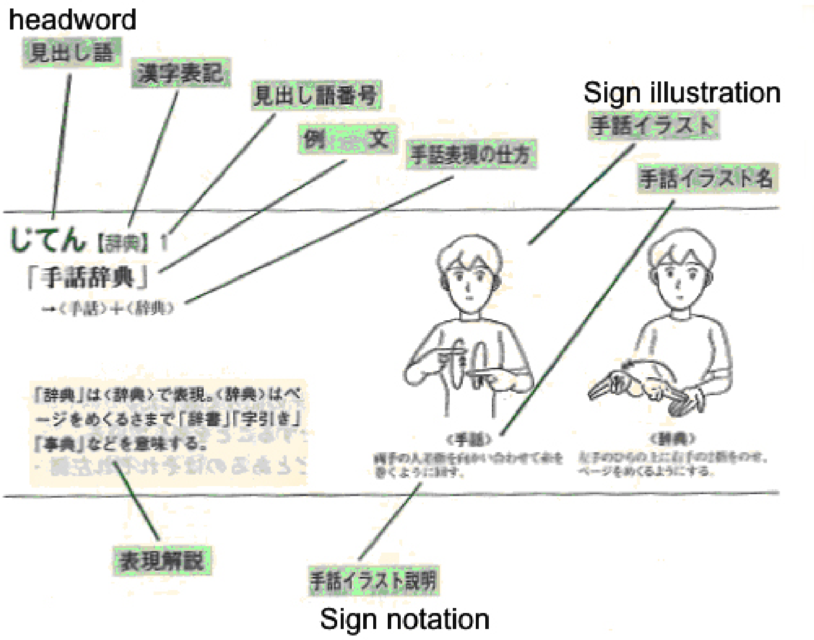


Figure 4. *Explanatory notes of Japanese-Japanese Sign Dictionary [Yonekawa, 1997]*

The authors scanned all pages of the dictionary and obtained notations via optical character reader. The obtained notations are analyzed through ChaSen, a morphological parser of the Japanese Language developed by NAIST Computer Linguistic Laboratory (Matsumoto et al, 2001). In morphological analysis, the authors applied medical corpus developed for natural language processing in medicine by the authors (Takemura and Ashida, 2001) to obtain proper anatomical nouns. The dictionary had 3249 noun or compound noun candidates including compound noun connected by the possessive particle "NO". To extract possible reference point, the authors selected 691 candidates including part of body manually. Analysis of the 691 candidates shows that the extracted compound nouns are consists of single prepositional modifier such as "HIDARI (left)" or "MIGI (right)", single or multiple nouns of body parts concatenated by possessive particle, and single postpositional modifier following a possessive particle such as "SHITA (under)" or "UCHIGAWA (inside of)" as shown in Fig 5, although possessive particles sometimes drop. Along this analysis, the authors divided selected compound noun candidates by the possessive particle into 142 compound nouns including part of body and 46 postpositional modifiers, and divide the 142 compound nouns into 4 prepositional modifiers and 85 nouns indicates part of body.

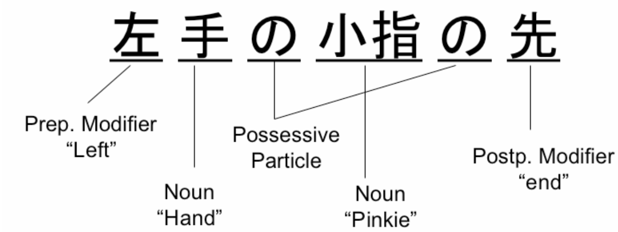


Figure 5. *Typical expression of reference points in sign notation.*

3.3 Applying the obtained reference points

The extracted 85 nouns obtained through semi-automatic morphological analysis may express same body part due to nature of Japanese Language, such as hiragana and Chinese character explanation for exactly same words. Manual investigation founds the nouns consists of 43 different body parts. The extracted 43 parts are successfully mapped on either a certain point of kinematical model or reference model of MPEG4 H-ANIM standard as Fig 6. Therefore, the proposed coordination system can be proportional to various avatars.

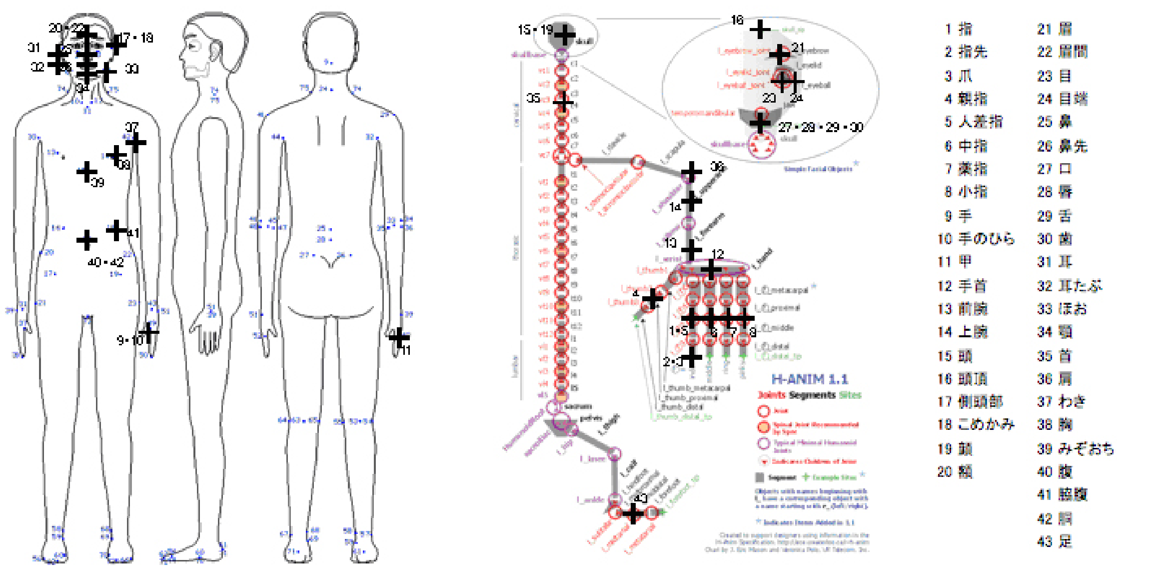


Figure 6. *Applying the reference points on H-ANIM avatar.*

4. DISCUSSION

The analysis of Japanese-Japanese sign language gives 85 body positions, which can be mapped on H-ANIM standard avatar. Thus the coordinate system may denote given sign motion in a certain manner.

On the other hand, the extracted modifiers and noun may produce 19975 compound nouns to show position or orientations of hands in disregard of multiple usages of nouns and drop of possessive particle. 321 out of the 19975 candidates appear 30919 times among 16640 notations in the dictionary in total. Regarding that the obtained notation including many errors of optical character reader, the proposed coordinate system can denote most of sign expressions appear in Japanese Sign Language Dictionary.

However, the further investigation found that the 321 compound nouns are also used to denote posture of hands. Thus, further morphological analysis regarding with verbs related to the 321 candidates may results in reducing number of nouns and modifiers to compose expression of position or orientation of hands. Additionally, further investigation of sign notations appears that some signs are expressed figuratively without using the compound words. To examine whether the obtained nouns are sufficient to show most of signs, to denote such signs using obtained reference points are required.

Although this paper proposed and mapped new coordinate system logically suitable for signs, sign language or human posture can be quite complicated. For example, to show “eye”, a signer can indicate his/her eye from any direction of eye if the fingertip of index is 0pt (or small amount) from targeting eye. Thus, to define motion of sign posture, to define dominant part of hand is indispensable. However, as the dominant part can be defined as nearest point to the reference point, or dependent on hand posture itself, automatic conversion from “raw” data into the proposed coordinate may be realized to a certain extent.

5. CONCLUSION

This paper proposed new coordinate system suitable for sign animation, and defined possible origin of coordinates through morphological analysis of Japanese Sign Language Dictionary. The investigation clears that the proposed coordinate can proportional to avatar and be applicable for most of sign expressions. The authors are performing further morphological analysis to refines possible reference points and to clear sufficiency of obtained reference points. Additionally, the authors are developing coordinate conversion to evaluate availability of the proposed coordinate system.

Acknowledgements: This research is partly supported by Grant-in-Aid for Young Scientists (A) (17680008) from The Ministry of Education, Culture, Sports, Science and Technology, Japan.

6. REFERENCES

P Ekman and W V Friesen (1978), *Facial action coding system: A technique for the measurement of facial movement*, Consulting Psychologists Press, Palo Alto.

T Hanke (2004) HamNoSys; representing sign language data in language resources and language processing contexts, *Proc. LREC 2004*, pp.1-6.

T Kurokawa (1992) Gesture Coding and a Gesture Dictionary for a Nonverbal Interface, *IEICE Transactions of Fundamentals of Electoronics*, 75, pp.112-121.

T Kuroda, K Sato and K Chihara (1998), S-TEL an avatar based sign language telecommunication system, *Int, J. Virtual Reality*, **3**, *4*, pp. 21-27.

T Kuroda, Y Tabata, A Goto, H Ikuta and M Murakami (2004), Consumer price data-glove for sign language recognition, *Proc. ICDVRAT*, pp.253-258.

T Kuroda, W Makari, Y Manabe and K Chihara (2001), Transformation of Signing Avatar to Fit User's Figure, *Moroccan J. of Control Computer Science and Signal Processing*, 3, *1*.

ISO/IEC (2001), Information Technology – Coding of Audio-Visual Objects- Part2: Visual, ISO/IEC 14496-2.

Y Matsumoto, A Kitauchi, T Yamashita, Y Hirano, H Matsuda, K Takaoka and M Asahara (2001) Moriphological Analysis System ChaSen Manual, Nara Institute of Science and Technology.

Y Nagashima and K Kanda (2001), Present State and Issue of Sign Linguistic Engineering, *Proc. HCI International*, **3**, pp. 387-391.

C Schmaling, T Hanke, (2004), HamNoSys 4.0, http://www.sign-lang.uni-hamburg.de-/Projekte/HamNoSys/HNS4.0/englisch/HNS4.pdf.

W C Stokoe, D. Casterline and C Croneberg (1965), *A Dictionary of American Sign Language on Linguistic Princeples*, Gallaudet University Press, Washington DC.

V Sutton (1998), The SignWriting Literacy Project, *Proc. Cognition AREA Conf.*

T Takemura and S Ashida (2001), A Concept of Medical Corpus for Natural Language Processing in Medicine, *Proc. the 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry*.

A Yonekawa (ed) (1997), *Japanese-Japanese Sign Language Dictionary*, Japanese Federation of the Deaf, Tokyo.

Session III. Therapy

Chair: Noomi Katz

Technological challenges and the Delft virtual reality exposure system

C A P G van der Mast

Department Man-Machine Interaction, Delft University of Technology,
Mekelweg 4, 2628CD Delft, THE NETHERLANDS

c.a.p.g.vandermast@tudelft.nl

mmi.tudelft.nl/~vrphobia

ABSTRACT

In this paper the architecture and use of the Delft VRET system is described. In this generic VRET system special emphasis is given to usability engineering of the user interface for the therapist. Results of controlled experiments with patients are summarized. The system is in regular use in a few clinics since 2005. New technological and functional challenges of VRET are presented. These challenges will lead to improvements of the system in the future. Agent support for the therapist and tele-VRET are the most promising challenges.

1. INTRODUCTION

Virtual reality exposure therapy (VRET) is an evolving technique that has attracted a lot of interdisciplinary research. VRET is the result of a close collaboration between researchers and practitioners of significantly different disciplines, among others, psychiatry, clinical psychology, psychotherapy, computer science, graphics design, human-computer interaction, and engineering. It has been shown that VRET is effective for patients with acrophobia, arachnophobia (spider phobia) and fear of flying (Emmelkamp et al, 2004). The effectiveness of VRET in other anxiety disorders like claustrophobia, fear of public speaking, fear of driving, posttraumatic stress disorder, and agoraphobia also holds promise for the future (Emmelkamp et al, 2004). At Delft University of Technology, in collaboration with department clinical psychology of University of Amsterdam, a generic system for treatment of phobia has been developed, taking into account a user-centered design process and specific human-computer interaction issues (Gunawan et al, 2004).

Traditional cognitive behavioural therapy has been taken as the main paradigm to be supported by technology in different ways, by providing interactive immersive worlds to "play" the treatment process in virtual reality instead of *in vivo*, as in the behavioural approach or by imagination, as in the cognitive framework. It has proved to be the case that patients are very sensitive to specific multimodal [i.e. more senses involved] stimuli in the virtual world (Krijn et al, 2004a). Medium-level resolution and graphics quality has proven sufficient in many cases to trigger the specific phobia-related reactions that are essential in exposure therapy. The effect of locomotion technique on fear is studied in (Schuemie et al, 2005). In one study (Krijn et al, 2004b), treatment using a standard head-mounted display (HMD) gave the same results for the treatment of acrophobia as a CAVE (computer automatic virtual environment) system providing advanced virtual reality systems. Of course this substitutability may be dependent on the specific type of disorder to be treated. It is proven in many studies that VRET can achieve the same results as traditional cognitive behavioural therapy, but will not outperform it (Emmelkamp 2005). But there are more aspects of cognitive behavioural therapy which are important besides exposure. Technology can also support the therapist in changing in real time to other synthetic worlds to be exposed to the patient, or in recording and replaying sessions in the virtual world for later analysis and planning the following session (Van der Mast et al, 2005).

Current VRET systems are mostly developed and used in laboratories where technical support is available. A few systems are available on the market, but evaluation of practical use on a larger scale has not yet been reported. To provide full support in the clinical roles, it is essential that VRET systems be usable in the clinic by several therapists of a team and without strong and expensive technical support. This usability is important to enhance the performance of the treatments on the one hand, and on the other hand we may expect benefits from other support functions beside the VR exposure technique itself. Interesting new technologies are available to extend a VRET system with new functions in order to measure and analyze

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

details of the treatment process for better understanding of diagnosis and treatment and for improving the efficiency of the therapist's work (Van der Mast et al, 2006).

The goal of this paper is to explore and discuss the technological and functional challenges to implement in VRET systems to be used in the clinic by regular therapists. On a scale that is following requirements given by society and economically interesting. We use the Delft VRET system as an example and model for these challenges.

2. THE ARCHITECTURE OF THE DELFT VRET SYSTEM

The functional architecture of the Delft VRET system (development started in 1999) is based on the main functions as depicted in Figure 1. This architecture is based on task analysis of therapists work by interviews and observations (Schuemie and Van der Mast, 2001; Schuemie, 2003).

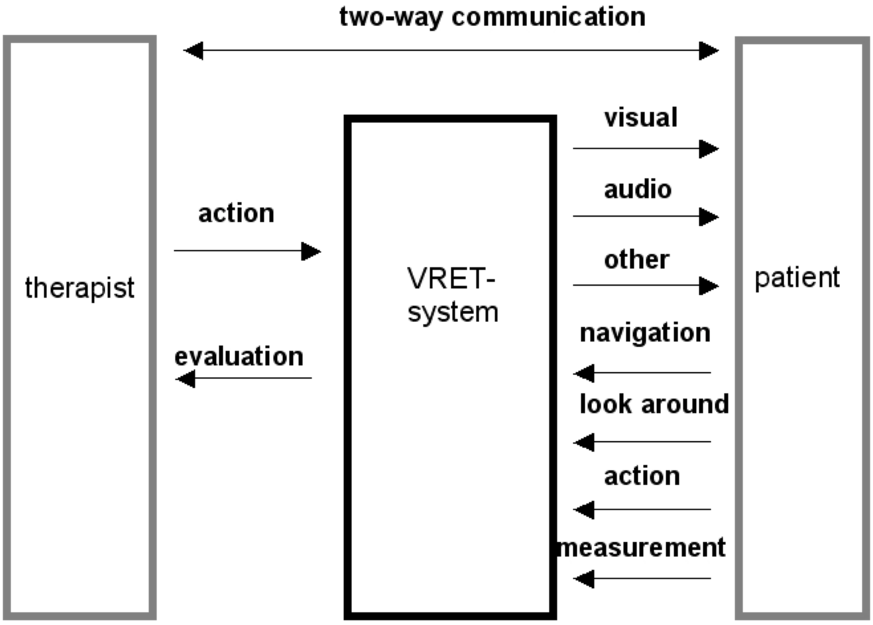


Figure 1. *The functional architecture of the Delft VRET system.The two-way communication of therapist and patient is direct if both are in the same room. An AV intercom connection is needed if both are not in the same room. The Delft VRET system is using two computers to be connected over the internet which supports tele-treatment.*

The two-way communication between therapist and patient is the base of the treatment. The VRET system is the intermediate tool to offer controlled stimuli to the patient. It is important that the therapist can control and evaluate the VRET system through a transparent user interface ("action" and "evaluation" in Figure 1). The user interface between patient and VRET system is in another way essential for the treatment. The patient should receive the visual, audio and other stimuli and she should be able to "navigate", "look around" and "act" in the virtual environment offered. The system should be provided with means to measure mental and physiological data of the patient. The therapist should be able to monitor and to supervise the behavior and experiences of the patient. In Figure 2 the therapist user interface is shown for a virtual world for treating fear of flying. The therapist can control the behavior of the plane (taxiing, taking off, weather conditions, pilot calls, turbulence) and she can follow how the patient is perceiving the virtual world by a live copy of patients view. She has also a second free view in the plane to look around to prepare next actions in the plane.

The technical architecture is based on using two independent computers, one for presenting the interactive worlds to the patient and one offering a control module to the therapist with the user interface (e.g. for fear of flying, see Figure 2). Both computers are connected using the internet protocol in such a way that the therapist is able to control the virtual worlds completely over the internet. This makes it possible to control sessions for tele-treatment. We tested this tele-medicine feature in the lab environment with collegues from the Universita Politecnica de Valencia.

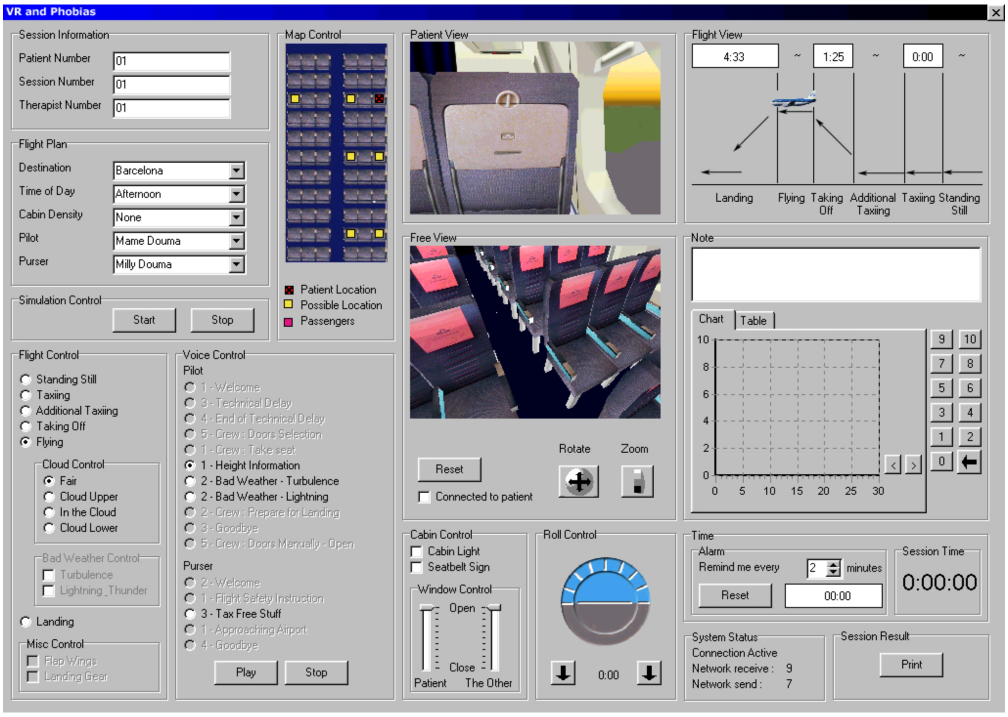


Figure 2. *The therapists user interface to monitor and control the airplane world for treating fear of flying. The upper image in the middle is the patients view. The lower image in the middle is the free view, including some controls for looking around by the therapist.*

3. USABILITY ENGINEERING

The evaluation of the therapist's UI is measured in terms of usability, i.e. effectivity, efficiency and satisfaction (Rosson and Carroll, 2002). We used a mediated evaluation which is a mix between analytical and empirical method, see Figure 3. The analytic evaluation is done early during the design process. The result of this analysis is used to motivate and develop materials for empirical evaluations. Heuristic evaluation as an inspection method was done. The empirical method is done by a user evaluation experiment. Because we have only a very limited number of professional therapists for treating phobias we added other persons to do the same treatment/job. From (Neerincx et al, 2001) we know that this may deliver valid results in usability studies for designing space stations. The therapists tested the system directly with a user/patient in the virtual world. Some specific tasks were given to the therapist to complete. Information was gathered, such as the observation protocol, performance time, errors and subjective evaluation. The subjective evaluation was acquainted by using usability questionnaires and interviews. To improve the therapist user interface of our fear of flying worlds we did usability engineering studies and found new functions to be included e.g. flight plan control, roll control and timer function. This is described in detail in (Gunawan et al, 2004).

Another outcome of our usability engineering approach was a world for treating social phobia. We implemented following the therapists requirements a market place (the Market in the city of Delft which is very well known to almost all inhabitants of the Netherlands) of which several parameters could be controlled online by the therapist during the session. Including weather conditions and number of walking avatars around the patient. How to design weather conditions, avatars, how to find the adjustable number of moving avatars, were all investigated using these usability engineering techniques, see (Van der Mast and Hooplot, 2006).

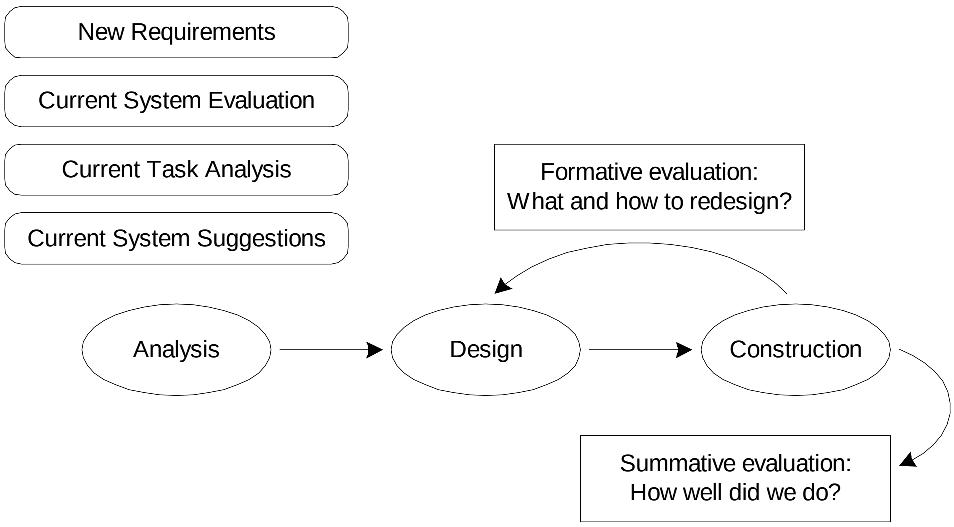


Figure 3. *Usability research methodology (Gunawan et al, 2004).*

4. THE RESULTS WITH THE SYSTEM

The first controlled experiment with the predecessor of the current Delft VRET system was done with *patients* who suffered from acrophobia ((Emmelkamp et al, 2001). This was a within-group design comparing in vivo treatment with VRET. Al patients were treated first with sessions of VRET followed by sessions of in vivo exposure. VRET was found to be as effective as exposure in vivo on all subjective measures. Conclusions were not yet firm because of the potential order effect.

In a second study (Emmelkamp et al, 2002) patients were assigned randomly to either VRET or exposure in vivo. The virtual worlds for acrophobia were built as an exact copy of the real worlds used in the condition in vivo: a shopping mall, a fire escape, and a roof garden (Figure 4a). VRET was found to be as effective as exposure in vivo on all measures. The improvements were maintained at 6-month follow-up.

(a) *(b)*

Figure 4. *(a) One of the worlds for treating acrophobia (Schuemie, 2003) and (b) the Delft VRET system in use at the clinic of PsyQ in the Hague.*

A third study (Krijn et al, 2004) focussed on the effectiveness of two different types of virtual environments, varying the degree of presence. A CAVE system was compared with a standard stereoscopic HMD. Patients were randomly assigned accross the two conditions. The virtual environments were implemented exactly the same being the shopping mall, the fire escape, and the roof garden used in the second study mentioned above on acrophobia. The results showed no differences in effectiveness between the two systems on all measures. As expected, presence was significantly higher in the CAVE, but this did *not* result in a more effective treatment.

In other studies and (Schuemie, 2003) this VRET system is used for fundamental research on the influence of patient navigation control modes walk-in-place, trackball, and gaze directed steering (Schuemie

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

et al, 2005), and the influence of patient versus therapist navigation control on patients presence. This last study showed no significant difference in presence. Krijn did (Krijn, 2006; Krijn et al, 2006) an extensive study on treatment of fear of flying.

In 2006 new worlds are being built for treatment of agoraphobia (van der Mast and Hooplot, 2006). Controlled experiments are not yet finished.

The Delft VRET system is being used from spring 2005 in the clinics of PsyQ in The Hague by several therapists for treating acrophobia and fear of flying of regular patients, and paid by their insurance companies (Figure 4b). The therapists report that the number of required VRET sessions is lower than with the traditional ways they are used to spend for treatment. This can mean an increase of efficiency. However, this should be confirmed by well designed controlled field experiments. In the summer 2006 the Delft VRET system is being used for regular treatment of fear of flying by VALK foundation in Leiden. VALK is specialized in fear of flying. They will also prepare controlled experiments.

5. NEW TECHNOLOGICAL CHALLENGES

We will discuss the following technical challenges. The order is arbitrary.

5.1 Personalizing the System

A VRET system may be used by different therapists from a clinic. Additionally, each therapist may change over time his or her preferences about using the system for some specific phobias. This gives a rationale for implementation the possibility to personalize the user interface and some of the main functions of the treatment process by the individual therapist and to store the applicable parameters. It is conceivable that this personalization could be extended to prepare for each patient an individual treatment procedure off-line, including some changes in the virtual worlds, specific for each patient to be treated. This kind of personalization is an important research goal sometimes referred to as “adaptive” user interfaces.

5.2 Automated Support for the Therapist

The first function of a VRET system is to offer an interactive virtual environment for the patient to experience the feelings that have to be worked on. But beyond that, the most promising challenge is to develop support functions for the therapist. By analyzing the treatment process and composing task models one can recognize and specify steps and modes in the treatment. In (Schuemie, 2003) examples of task models (treatment procedures, see Figure 5) are described. They are based on observation of sessions. New models have to be developed describing the task in terms of treatment steps and specific aspects and levels of the disorder.

If it is possible to describe a treatment session in terms of steps to be taken under supervision by the therapist it may be possible to develop an electronic agent which provides advice including some rationales to the therapist about the next possible step(s) in the actual context of the treatment. It would also seem interesting to provide a planning mode to the therapist to specify some sequential steps for a session just before it starts. The general goal is to provide extra explicit knowledge to the therapist about the progress of the treatment. The agent can obtain its information from built-in procedures which may be adjusted by the therapist and by measurements of the patients physiological condition, e.g. heart rate. It would be most useful to construct a learning electronic agent that could learn from an experienced therapist. A junior therapist could use this “smart” agent to give better treatment in non-critical sessions, under the responsibility and supervision of an experienced therapist. It would be possible to teach such agents to give good advice by analyzing individual treatment patterns in specific clinical cases. The advice might propose the next procedure step or the next navigation or modification of the virtual world to control the level of fear. Even more measurements can be done by voice-emotion recognition since these can indicate levels of stress, fear, and other emotions. This voice-emotion recognition may replace the standard asking for SUDs (selective units of discomfort).

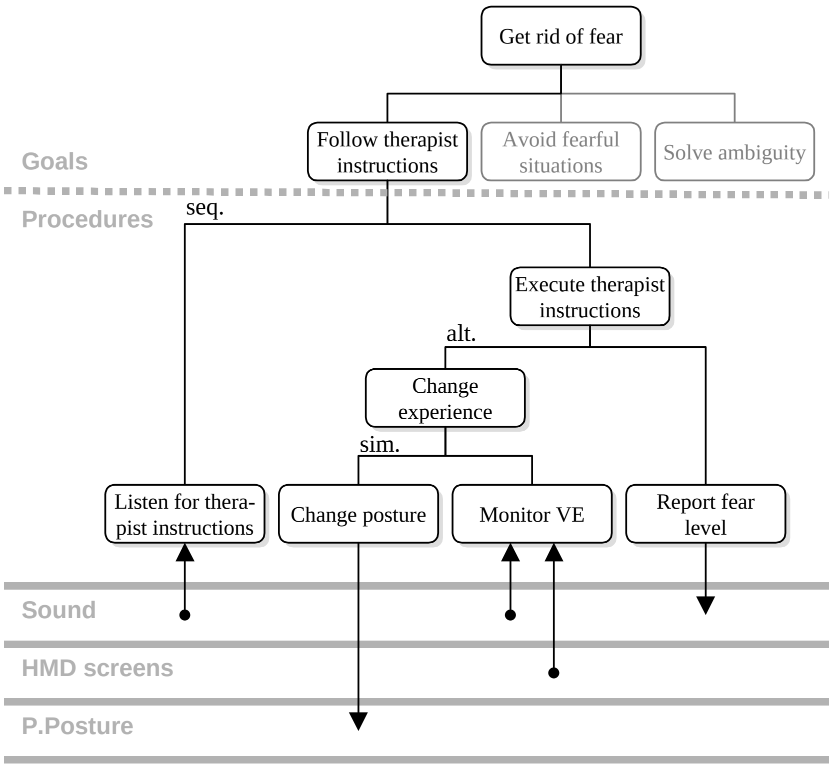


Figure 5. *'Follow therapist instructions' goal decomposition (Schuemie, 2003). Sound, HMD screens and P. Posture are the related communication modes (senses) between patient-machine and patient-therapist. VE is virtual environment.*

As an additional form of support may be considered procedures of computer-supported self-treatment by the patient. The therapist should be able to specify the procedures and constraints of these modules for self-treatment. The therapist could individualize homework scenarios for the patients. Individualization of virtual environments is addressed within a currently ongoing EMMA project (Botella, 2006).

5.3 Computer-based training

A completely different challenge is the construction of a VRET system for computer-based training of junior therapists. This could be done using simulated or real patients or recorded sessions. The learner could be trained how to use the system and how to treat different types of disorders. Simulated patients could include a combination of computer-generated patients and normal people requested to simulate. Computer-generated patients could model non-visible changes happening in real patients (e.g. physiological changes).

5.4 Tele-Care

Because the therapists report to us that the habituation process is going slowly with relatively less interaction between therapist and patient, they express to have time to treat more patients simultaneously, given a good intercom system for personal communication with each patient simultaneously. It is both a technical and an organizational challenge to develop a system for tele-treatment of mental disorders using VRET over the internet. The most serious challenge is to have a ratio *therapist : patient* of more than 1:1. It should be possible to develop a system and a therapists user interface to allow the provision of treatments to more than one patient at the same time, in different rooms in the same clinic or in different clinics. If one senior and one junior therapist could treat more than two patients simultaneously, the ratio will improve. Some experience with tele-treatment of agoraphobia without VR has previously been reported (Lange at al, 2000). In a more general project on tele-care the possibilities of agent support for tele-care at home has been investigated (Riva et al, 2004).

6. CONCLUSIONS

We have seen that several interesting technological challenges are on the horizon. But we must remain aware that we need fundamental research on how new technologies can improve the very personal treatment process supervised by the therapist. This research must be demand-driven by the therapists, and not pushed by technology. We are just in the early stages of some very interesting developments. They will improve our insight in how treatment can be given in the most effective way and how treatment can be deployed on a large scale more efficiently than with the current means. VRET may play an important role in these developments.

In our view, an emerging scenario could characterize the future clinical setting: old (and functional) practices could be integrated and enhanced through new (and promising) media such as VR. The most promising challenges are agent support for the therapist and tele-VRET. The Delft VRET system will be used to work on these challenges.

Acknowledgements: The author likes to thank Paul Emmelkamp, Martijn Schuemie, Merel Krijn, and many participating master-students for their essential contribution of the Delft VRET system.

7. REFERENCES

C Botella, H Villa, A Garcia-Palacios, R M Banos, C Perpina, M Alcaniz (2004), Clinically Significant Virtual Environments for the treatment of Panic Disorder and Agoraphobia. *Cyberpsychology & Behaviour*, **7**, 5, pp. 527-535.

C Botella (2006), Clinical issues in the application of virtual reality to treatment of PTSD, *Proc. NATO workshop Novel approaches to the diagnosis and treatment of posttraumatic stress disorder.* IOS Press, Amsterdam, pp. 183-195.

P M G Emmelkamp, M Bruynzeel, L Drost, C A P G van der Mast (2001),Virtual Reality Treatment in Acrophobia: A Comparison with Exposure in Vivo, *Cyberpsychology and Behavior*, **4**, 3, pp.335-341

P M G Emmelkamp, M Krijn, A M Hulsbosch, S de Vries, M J Schuemie and C A P G van der Mast (2002), Virtual Reality Treatment versus exposure in vivo: A Comparative Evaluation in Acrophobia, *Behaviour Research & Therapy*, **40**, 5, pp.509-516.

P M G Emmelkamp, M Krijn, R P Olafsson, R Biemond (2004), Virtual reality exposure therapy of anxiety disorders: A review, *Clinical Psychology Review*, **24**, pp. 259-281.

P M G Emmelkamp (2005), Technological innovations in clinical assessment and psychotherapy, *Psychotherapy & Psychosomatics*, **74**, pp. 336-343. (2005)

L T Gunawan, C A P G van der Mast, M Krijn, P M G Emmelkamp, M A Neerincx (2004), Usability of therapist's user interface in virtual reality exposure therapy for fear of flying, *Proc. of Euromedia 2004 Conference*, Hasselt Belgium, pp. 125-132.

M Krijn (2006), *Virtual Reality and Specific Phobias --Welcome to the real world*, PhD thesis, University of Amsterdam. (available from http://mmi.tudelft.nl/~vrphobia)

M Krijn, P M G Emmelkamp, R P Olafsson, M Bouwman, L van Gerwen, M J Schuemie and C A P G van der Mast, (2006), Fear of flying: A comparative evaluation of Virtual Reality Exposure Therapy, Cognitive Behavioural Methods and Bibliotherapy, (submitted).

M Krijn, P M G Emmelkamp, R P Olafsson, R Biemond (2004a), Virtual reality exposure therapy of anxiety disorders: A review, *Clinical Psychology Review* **24** pp. 259-281.

M Krijn, P M G Emmelkamp, R Biemond, C de Wilde de Ligny, M J Schuemie and C A P G van der Mast (2004b), Treatment of acrophobia in Virtual Reality; the role of immersion and presence, *Behaviour Research and Therapy*, **42**, 2:229-239.

A Lange, B A L Schrieken, J Q R van de Ven, B Bredeweg and P M G Emmelkamp (2000), INTHERAPY: The effects of a short protocolled treatment of post-traumatic stress and pathological grief through the internet. *Behavioral* & *Cognitive Psychotherapy*, **28**, 2, 103-120.

C van der Mast, S Popovic, D Lam, G Castelnuovo, P Kral, and Z Mihajlovic (2006), Technological challenges in the use of Virtual Reality Exposure Therapy, *Proc. of the NATO Advanced Research Workshop on Novel approaches to the diagnosis and treatment of posttraumatic stress disorder*, Amsterdam IOS Press, pp. 286-295.

C van der Mast and F Hooplot (2006), Design and evaluation of a VRET system for agoraphobia, *Proc. of Euromedia 2006 Conference*, Athens Greece, pp. 77-84.

M A Neerincx, M Ruijsendaal and M Wolff (2001), Usability Engineering Guide for Integrated Operation Support in Space Station Payloads, *International Journal of Cognitive Ergonomics*, **5,** 3, 187-198.

G. Riva, F. Morganti and M. Villamira (2004), Immersive virtual Telepresence: virtual reality meets eHealth, *Cybertherapy* (G. Riva et al. Eds) , IOS Press Studies in health technology and informatics, **99**, pp. 255-270.

M B Rosson and J M Carroll (2001), *Usability Engineering - scenario-based development of human-computer interaction*, Morgan Kaufmann Series in Interactive Technologies, Academic Press, San Fransisco.

M J Schuemie, P van der Straaten, M Krijn, C A P G van der Mast (2001), Research on Presence in VR: a Survey, *Cyberpsychology and Behavior*, **4**, 2, April 2001, pp.183-202.

M J Schuemie, and C A P G van der Mast (2001), VR Testbed Configuration for Phobia Treatment Research, *Proc. of Euromedia 2001 Conference*, Valencia, Spain, pp. 200-204.

M J Schuemie, C A P G van der Mast, M Krijn and P M G Emmelkamp (2002), Exploratory Design and Evaluation of a User Interface for Virtual Reality Exposure Therapy , in: J.D. Westwood, H.M. Hoffman, R.A. Robb, D. Stredney (Eds.), *Medicine Meets Virtual Reality 02/10*, IOS Press, Amsterdam, pp.468-474

M J Schuemie (2003), *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*, Ph.D thesis, Delft University of Technology. (available from http://mmi.tudelft.nl/~vrphobia)

M J Schuemie, B Abel, C A P G van der Mast, M Krijn and P M G Emmelkamp (2005), The effect of locomotion technique on presence, fear and usability in a virtual environment, *Proc. of Euromedia 2005 Conference*, Toulouse France, pp. 129-135.

Evaluation of virtual reality therapy in augmenting the physical and cognitive rehabilitation of war veterans

B K Wiederhold[1,2] and M D Wiederhold[2]

[1]Interactive Media Institute, 6160 Cornerstone Court, Suite 155, San Diego, CA 92121

[2]Virtual Reality Medical Center, 6160 Cornerstone Court, Suite 155, San Diego, CA 92121

[1]*bwiederhold@vrphobia.com*

[1]*www.interactivemediainstitute.com,* [2]*www.vrphobia.com*

ABSTRACT

War veterans with neuromuscularskeletal injury often require significant treatment and rehabilitation, straining health care resources. In a study funded by the Office of Naval Research (ONR) and the Defense Advanced Research Projects Agency (DARPA), the Virtual Reality Medical Center (VRMC) is applying virtual reality therapy to injured military personnel at the Naval Medical Center San Diego (NMCSD). The goal of this study is to investigate whether augmenting traditional rehabilitation with VR (in this case, off-the-shelf interactive video games) will enable a more rapid and complete rehabilitation. Because VR is interactive, and encourages patients to use their entire body to reach goals in the game, it is conceivable that it will make rehabilitation sessions more comfortable and entertaining. Participants consist of 20 veterans with upper arm and shoulder injuries (rotator cuff tear, shoulder impingement, bursitis) or amputation. The participants are divided into two groups (n=10): an experimental group, which receives traditional rehabilitation augmented by virtual reality therapy, and a control group, which undergoes traditional rehabilitation. Participants will complete ten treatment sessions in their respective condition. Though the study has not been completed, preliminary results based on subjective questionnaires and functional capacity indicate that the experimental condition may elicit increased heart rate and respiration. Participants in this group also seem to enjoy the music and interaction made possible through VR. These results suggest that VR may enhance the rehabilitation process, creating a more effective form of treatment. Long-term benefits of this form of treatment may include improved treatment time and reduced drop out rates, therefore reducing the costs of rehabilitation.

1. INTRODUCTION

Rehabilitation is designed to restore a patient's physical, sensory, and/or mental capabilities lost due to injury, illness, or disease. *Cognitive* rehabilitation is a structured set of therapeutic activities designed to retrain an individual's ability to think, use judgment, and make decisions. *Physical* rehabilitation is the process of restoring a patient's lost function through hands-on treatment, exercise, and patient education. Cognitive and physical rehabilitation are among the largest sectors of healthcare costs in the United States. Total treatment expenditures are growing rapidly due to the aging of the United States population, as well as to higher rates of battle-related injury.

As of August 2005, it was estimated that over 12,500 U.S. warfighters had been wounded in Iraq. Because of advances in protective armor and battlefield medical care, more military personnel are surviving their wounds. This means that more people with amputated limbs, traumatic brain injury, and other injuries are returning to the U.S. for rehabilitation and treatment. The long term care of thousands of wounded veterans promises to be a large expense (Glasser, 2005).

This tremendous cost burden to the healthcare system generates an immediate need for new clinical approaches to both improve the efficacy of physical and cognitive rehabilitation and to reduce the total cost of such treatment. Recent improvements in computer and sensor technology now make it possible to develop portable home telerehabilitation systems that have the potential to dramatically improve rehabilitation outcomes for neurological and musculoskeletal injuries, while reducing overall rehabilitation costs by decreasing the need for in-person treatment (Hoenig, et al. 2006; Lewis, Deutsch, & Burdea, 2006; Sugarman, Dayan, Weisel-Eichler, & Tiran, 2006).

Aside from the obvious hurdle of expense, rehabilitation treatments are often less effective because they are unpleasant or monotonous. Duncan et al. (2002) found that compliance and adherence to rehabilitation therapy and guidelines are associated with improved treatment outcomes. While the focus of their research was stroke patients, it is not difficult to see how this principle would apply to other populations. Virtual Reality (VR) has the potential to help achieve the compliance necessary for effective rehabilitation. VR refers to computer-generated 3-D interactive synthetic environments, examples of which include video games. By increasing the motivation and enthusiasm for a patient's compliance with their physical rehabilitation therapy, VR becomes an advantageous addition to traditional physical therapy for patients.

For example, in a study with seven post-stroke elderly patients undergoing physical rehabilitation therapy for a weak arm, patients interacted with a virtual environment (VE) using a Sony PlayStation II EyeToy™. Researchers encouraged patients to use their weak arms when interacting with the games in the VE. By the end of the experiment, two chronic stroke patients had succeeded in using their weak arm to clean the left side of a virtual window. All of the patients reported that they had enjoyed the experience (Rand, Kizony, Weiss, 2004).

Other studies have supported this result. In a research study by Sveistrup et al. (2003), individuals presenting frozen shoulder due to musculoskeletal pathology went through exercise-training sessions, three times a week for six weeks, under three conditions: 1) conventional stretching, 2) traditional exercise, and 3) virtual reality exercise. Results indicated that virtual applications for rehabilitation increased the amount of interaction between patients and their environment. In 2001, Aaron, Rose, Janesen, and Hentz used a glove-based VR system for rehabilitation in patients with carpal tunnel syndrome and other hand-related injuries. Results showed that participants complied better with their rehabilitation program when they used a glove-based VR system than when they used traditional physical therapy alone. Other studies with stroke patients using VR therapy have shown improvements in arm movement and upper extremity functions (Piron, et al., 2003; Piron, et al., 2005; Holden, Dyar, Schwamm, & Bizzi, 2005). In some patients, improvements in arm movement have also corresponded with improved scores on measurements of independence of ADL (Piron, et al., 2003; Piron, et al., 2005). VR has also been used to retrain upper limb motor movements, e.g., pouring water, in patients with acquired brain injury (Holden, Bettwiler, Dyar, Niemann, & Bizzi, 2001).

The advantages of VR extend to cognitive rehabilitation as well. It can be used for vocational training and as a way to train cognitive tasks in brain damaged patients (Rose, Brooks, & Rizzo, 2005). VR research in the field of motor rehabilitation includes applications for stroke, acquired brain injury, Parkinson’s disease, orthopedic rehabilitation, balance training, wheelchair mobility, and training in functional activities of daily living (Holden, 2005). Research in VR applications for stroke patients in particular have made considerable progress; fMRI studies of VR used in conjunction with traditional physical therapy in stroke patients have resulted in neuroplastic changes in the brain and corresponding improvements in motor functions (Jang, et al., 2005). A study by Hofmann et al. (2003) asked ten patients suffering from mild to moderate Alzheimer's disease to move through a virtual environment that simulated pictures of the patient's typical surroundings, home or usual shopping route, and photographs of the patient at an earlier age. Results indicated that after three weeks of training, patients could perform the task more quickly and needed less help navigating the environment. In addition, eight out of the ten patients made fewer mistakes.

Furthermore, patients with phantom limb pain, a frequently debilitating phenomenon among amputees, have responded to the application of virtual reality technology. Professor Jonathan Cole of the Poole Hospital, University of Bournesmouth, and the University of Southampton in the United Kingdom, reported on the use of synthetic environments for the alleviation of phantom limb pain (Gallagher, 2004). The enhancement of missing-limb perception through the use of VR technology has been equally effective in alleviating phantom limb pain, while at the same time permitting a much greater perceptual range of absent-limb motion. Currently, a number of experimental studies are exploring virtual reality as an assessment tool for patients in the healthcare system. Virtual Reality and other advanced technologies may serve as an objectifying method to track performance improvements over the course of therapy. In light of the high number of amputees in the veteran population in the United States, this application might be especially relevant to military needs.

2. VIRTUAL REHABILITATION

2.1 *Rationale*

There are many reasons why VR applications are so effective for rehabilitation. First, VR is an interactive, experiential medium. In the same way that children and teenagers intuitively grasp computers, VR users become directly engaged with the effects of the virtual experience. In addition, VR creates a safe setting where patients can explore and act without feeling threatened (Riva, 2005). Patients can make mistakes without fear of dangerous, real, or humiliating consequences (Standen & Brown, 2005). Moreover, unlike human trainers, computers are infinitely patient and consistent (Standen & Brown, 2005). In cognitive rehabilitation, VR can be manipulated in ways that the real world cannot. For example, VR can convey rules and abstract concepts without the use of language or symbols for patients with little or no grasp of language. Indeed, VR has been used to support individuals with intellectual disabilities by training them in many skills of independent living, such as grocery shopping, food preparation, orientation, road crossing, and vocational training (Standen & Brown, 2005).

In motor rehabilitation, there are essentially three major advantages that VR offers over traditional therapy alone. First, VR creates a safe, controlled environment for repetitive practice, and repetitive practice is crucial in learning motor tasks. Second, VR provides immediate, real-time feedback about performance. Finally, because of its interactive nature, VR can increase motivation by making the experience fun (Holden 2005).

The current study utilizes an experimental study design in which one subset of patients (the control group) will receive traditional rehabilitation treatment, while another subset of patients (the experimental group) will receive traditional rehabilitation treatment *augmented by virtual reality therapy.*

Off-the-shelf *kinesthetic* video games, such as Sony's Eyetoy™, Dance Dance Revolution and Taiko Drum Master, will be utilized as virtual reality therapy. Kinesthetic video games focus on the player's body movements for input and control instead of a traditional joystick or controller. (See Figures 1 and 2.) For example, Sony's Eye Toy contains soccer, window washing, and dancing scenarios that require the user to use arm and body movements to block balls, eliminate spots on a window, or complete specific dance moves. Dance Dance Revolution utilizes a game pad and asks users to step on specific areas of the pad at certain times. Taiko Drum Master is played by hitting the drum peripheral in time with notes traveling across the screen. Sensors in the drum's surface and rim record the accuracy of each hit. Tasks increase in difficulty from single notes, to drum rolls, to different rhythms and speeds.

The body and arm movements required to play these games could encourage improvement in balance, strength and range of motion. Due to the nature of kinesthetic video games, it is conceivable to use them for physical therapy in order to make treatment sessions more comfortable and entertaining, in turn making them less tedious.


Figures 1 and 2. *Screen shots from off-the-shelf video games utilizing Sony's Eyetoy™. Kinesthetic video games such as these focus on the player's body movements for input and control instead of a traditional joystick or controller.*

The progress of patients in the control group and the experimental group will be measured using psychometric tests, physiological monitoring, and standardized physical therapy assessment tools. The outcomes of the control group will be compared to those of the experimental group. Based on existing research, the following hypotheses are posited for the current study:

- *H1.* As compared to the control group, the experimental group will exhibit greater improvements in physical and psychological measures.
- *H2.* As compared to the control group, the experimental group will exhibit greater improvements in physical and psychological measures.

2.2 Methodology

2.2.1 Participants. The present study recruited 20 participants with arm amputations or upper arm and shoulder injuries (e.g. rotator cuff tear, shoulder impingement, bursitis). Participants were mixed in age, sex, and education, and were randomly assigned to the control group (n=10) or the experimental group (n=10). Each participant will complete ten sessions of treatment, taking part in either traditional rehabilitation activities (control condition) or traditional activities with the addition of VR (experimental condition). Participants endorsing current suicidal or homicidal ideation, meet diagnostic criteria for bipolar disorder, schizophrenia, or current substance dependence using the MINI were excluded from the study.

2.2.2 Materials. We are using off-the-shelf inexpensive video games to augment physical therapy and rehabilitation exercises. These inexpensive, interactive and entertaining tools are simple, but can serve as a platform to ask basic questions concerning patient and medical staff acceptance of this new format. We are conducting a clinical evaluation using this kinesthetic video game intervention to better understand the requirements necessary to develop and integrate a more tailored program that more specifically addresses the needs of patients undergoing therapy from injuries encountered in wartime.

2.2.3 Measures. Pre-test questionnaires include the Virtual Environment Pre-Treatment Questionnaire, the Beck Depression Inventory II (BDI-II), the Physical Therapy Evaluation Report, and the Overall Health Status Assessment. Participants will also complete a pre-participation physical evaluation administered by their physician. After therapy, the self-report questionnaires are completed again (BDI-II and Overall Health Status Assessment). In addition, participants fill out post-test questionnaires: an exit questionnaire, Virtual Environment Post-Treatment Questionnaire, and usability questionnaire. An evaluation will also be administered by the physical therapist. Follow up may be by phone or in person.

2.2.4 VR Scenario and Testing Procedure. Once participants complete all relevant pre-treatment consent forms and questionnaires, they are given an educational session to orient them to the equipment and methods used in the study. After the appropriate video game is chosen, participants were randomly assigned to one of two groups: the VR intervention group (those using the kinesthetic video games in addition to traditional physical therapy) or the control group (those experiencing traditional physical therapy treatment). Participants in both groups undergo physical therapy under direct supervision of their physical therapist and a research assistant who records subjective assessments.

In the experimental condition, participants complete traditional rehabilitation exercises and movements while using the kinesthetic video games. In these games, tasks progressively increase in difficulty. Participants are encouraged to attempt to earn a high score during each session with the games. At each session, participants' progress is monitored using psychometric tests, physiological monitoring, and standardized physical therapy assessment tools. After a full course of treatment, each participant completes the post-treatment questionnaires.

3. RESULTS

Though the current study is still underway, and therefore, firm results are not yet available, trends have begun to emerge during testing. As the study progresses, it appears that those in the experimental group (VR augmented rehabilitation) have an increased heart rate and respiratory rate during each session. In addition, motivation may be improved with the use of VR as subjective reports reveal that the participants enjoy the music and interaction provided by the VR.

4. POTENTIAL IMPLICATIONS

Due to the successful background in using simulation technology for the assessment and treatment of neuromuscular disorders, neural cognitive disorders and musculoskeletal injuries, there is a significant opportunity to apply these new treatment paradigms to those injured in the Iraq War. If the current study demonstrates, as hypothesized, that virtual reality therapy augments rehabilitation progress, the next step will be the development and application of more advanced rehabilitation technologies and strategies. That is, the current study has the potential to catalyze improvement and change within clinical rehabilitation at large.

While these technologies will have immediate benefit for injured military personnel, their development will also serve to catalyze improvement and change within clinical rehabilitation at large. VR may indeed help create a more enjoyable and effective method of rehabilitating patients with war-induced injuries than the current paradigm.

Acknowledgements: The authors would like to acknowledge the Office of Naval Research (ONR) and the Defense Advanced Research Projects Agency (DARPA) for supporting this project.

5. REFERENCES

D H Aaron, B W Rose, C Janesen, V R Hentz, (2001). Pilot study of the Greenleaf Telerehabilitation System. Greenleaf Medical Telerehabilitation Executive Summary.

P W Duncan, R D Horner, D M Reker, G P Samsa, H Hoenig, B Hamilton, B J LaClair, T K Dudley (2002), Adherence to postacute rehabilitation guidelines is associated with functional recovery in stroke, *Stroke, 33*(1), pp. 167-77.

S Gallagher (2004), Nailing the lie: An interview with Jonathan Cole, *Journal of Consciousness Studies, 11*(2), pp. 3-21.

R J Glasser (2005), A ware of disabilities: Iraq's hidden costs are coming home, *Harper's Magazine, http://uslaboragainstwar.org/article.php?id=8862*.

H Hoenig, J A Sanford, T Butterfield, P C Griffiths, P Richardson and K Hargraves (2006), Development of a teletechnology protocol for in-home rehabilitation, *Journal of Rehabilitation Research and Development, 43*(2), pp. 287-298.

M Hofmann, A Rosler, W Schwarz, F Müller-Spahn, K Kräuchi, C Hock, et al (2003), Interactive computer-training as a therapeutic tool in Alzheimer's disease, *Comprehensive Psychiatry, 44*, pp. 213-219.

M K Holden, A Bettwiler, T Dyar, G Niemann, E Bizzi (2001), Retraining movement in patients with acquired brain injury using a virtual environment, *Stud Health Technol Inform, 81*, pp.192-8.

M K Holden, (2005), Virtual environments for motor rehabilitation: review, *CyberPsychology & Behavior, 8*(3), pp.187-211.

M K Holden, TA Dyar, L Schwamm, E Bizzi (2005), Virtual-environment-based telerehabilitation in patients with stroke, *Presence, 14*(2), pp. 214-233.

S H Jang, S H You, M Hallett, Y W Cho, C M Park, S H Cho, HY Lee, T H Kim (2005), Cortical reorganization and associated functional motor recovery after virtual reality in patients with chronic stroke: an experimenter-blind preliminary study, *Arch Phys Med Rehabil, 86*, pp. 2218-23.

J A Lewis, J E Deutsch, G Burdea (2006), Usability of the remote console for virtual reality telerehabilitation: formative evaluation, *CyberPsychology and Behavior, 9*(2), pp. 42-147.

L Piron, P Tonin, AM Atzori, C Zucconi, C Massaro, E Trivello, M Dam (2003), The augmented-feedback rehabilitation technique facilitates the arm motor recovery in patients after a recent stroke, *Stud Health Technol Inform, 94*, pp. 265-267.

L Piron, P Tonin, F Piccione, V Iaia, E Trivello, M Dam (2005), Virtual environment training therapy for arm motor rehabilitation, *Presence, 14*(6), pp. 732-740.

D Rand, R Kizony, P L Weiss (2004), Virtual reality rehabilitation for all: Vivid GX versus Sony PlayStation II EyeToy, *Proceedings of the 5th International Conference on Disabilities, Virtual Reality, & Associated Technologies*, Oxford, UK, Sept. 18-20, 2004, pp. 87–94.

G Riva G (2005), Virtual Reality in Psychotherapy: Review. *CyberPsychology & Behavior, 8*(3) pp. 220-230.

F D Rose, B M Brooks, A A Rizzo (2005), Virtual Reality in Brain Damage Rehabilitation: Review, *CyberPsychology & Behavior, 8*(3), pp. 241-262.

P J Standen, D J Brown (2005), Virtual Reality in the Rehabilitation of People with Intellectual Disabilities: Review, *CyberPsychology & Behavior, 8*(3), pp. 272-282.

H Sugarman, E Dayan, A Weisel-Eichler, J Tiran (2006), The Jerusalem TeleRehabilitation System, a new low-cost, haptic rehabilitation approach, *CyberPsychology and Behavior, 9*(2), pp. 178-182.

H Sveistrup, J McComas, M Thornton, S Marshall, H Finestone, A McCormick, et al (2003), Experimental studies of virtual reality-delivered compared to conventional exercise programs for rehabilitation, *Cyberpsychology & Behavior, 6,* pp. 245-249.

Do we need high-scale flexibility in virtual therapies?

Z Geiszt[1], M E Kamson[1], C Sik Lányi[1] and J A Stark[2]

[1]Department of Image Processing and Neurocomputing, University of Pannonia
8200 Veszprém, Egyetem u. 10. Hungary

[2]Department of Linguistics and Communication Research, Austrian Academy of Sciences
Kegelgasse 27/1, AU – Vienna

geiszt@vision.vein.hu, *moremi_kamson@yahoo.com*, *lanyi@almos.vein.hu*,
jacqueline.stark@oeaw.ac.at

[1]*www.knt.vein.hu*, [2]*www.oeaw.ac.at*

ABSTRACT

Virtual reality (VR) offers a wide range of applications in the field of cognitive neuropsychology both in diagnosing cognitive deficits and in treating them. An optimal diagnostic method is the on-field test, which provides an opportunity to apply VR-based simulations. VR is also a useful tool for skill-building and training by setting up a virtual setting, which imitates the real environment including the attributes to be trained. Moreover, it provides a graded approach to problem-solving and the feeling of safety and it excludes the negative elements which are detrimental to the learning process. To further extend the effectiveness of VR applications it is necessary to refine VR environments and adapt them according to the specific needs of selected target groups and the real-time control of virtual events. Following the principle of flexibility we prepared two virtual environments: 1) an Adjustable Virtual Classroom (AVC) for the treatment of fear of public speaking in a primary school task-solving setting, and 2) a Virtual Therapy Room (VTR) designed for use with aphasic clients. Due to their flexible nature, a large number of elements can be customised in both of these settings including spatial organisation, textures, audio materials and also the tasks to be solved. The real-time control over the virtual avatars by the supervisor, i.e. therapist to guide the social interactions in the virtual world also allows him/her to follow-up on the user's reactions and therapy performance. By focusing on the details of the therapy room, we would like to demonstrate the relevance of the flexibility of the software in the development of innovative therapy solutions for aphasic clients.

1. INTRODUCTION

Virtual reality offers a wide range of applications in the field of cognitive rehabilitative medicine. The research area we are interested in is rehabilitation in the field of cognitive neuropsychology and in particular language therapy for persons with aphasia. In this area, the most significant application of VR is in the therapy environment as a tool for intensive structured practice for the patient. However, it also has its merits in diagnosing neuropsychological deficits. As stated by Pugnetti and colleagues (1998): "VR will probably change some of the diagnostic procedures that are still used after decades have passed from their introduction". For this reason, various neuropsychological test procedures should be adapted according to innovations in the field of VR. On the other hand, Shallice (1991) considers the on-field tests to be the best tool for diagnostic purposes, because they provide valid patient information regarding the use of VR-based simulations.

The issue of cost/benefit ratio is the first to address when considering applying VR solutions in this field. This has actually been done, at least with reference to neuropsychologists' being willing to open up to this new technology. They are well aware of the fact that VR technology does not consume an immeasurable amount of money, since most of the applications suitable for their purposes are based on low-cost technology and off-the shelf components (cf. Pugnetti et al, 1998).

Botella and colleagues (1998a) also view VR as a "research laboratory setting for clinical psychology", where emotions, thoughts, and behaviour of patients can be studied through imitating a situation, an atmosphere in which the persons reactions are available for monitoring. Obviously, VR is not suited for the treatment of every type of cognitive deficit or mental disorder. However, for example in the case of phobias, the patients can be exposed to the virtual version of their fears and the clients can profit from this exposure.

As the well-known application of VR in the field of rehabilitative medicine is concerned, namely a therapeutic environment allowing direct practice for improvement of skills, numerous publications report positive results, e.g.: in the treatment of acrophobia, fear of flying and claustrophobia (Botella et al., 1998b; North and North, 1996, Rothbaum et al., 1995; Laky and Sik Lányi, 2003; Sik Lányi et al., 2004, etc.).

The environmental enrichment theory – a well known theory in neuroscience - also supports the importance of situational exercises, which is clearly associated with VR-based simulation. With regard to studies with nonhuman subjects, the term 'enriched environment' refers to an environment with numerous playthings for the animals and less food intake inducing competition among the animals. Renner and Rosenzweig reported (1987) that when laboratory animals were exposed to an enriched environment, they showed a significant improvement in their behavioural patterns. They were superior to their non-enriched counterparts with respect to complex learning and problem solving tasks. Their advantage increased with an increase in the complexity of the problem to be solved. It was observed that enrichment induced changes were evident in several parts of the brain as demonstrated by an increased number of neuron combinations, a heavier cortex, which is deeper and greater in surface area. Kempermann and colleagues (1997) indicated that these factors contribute to increased performance.

Similar result have been found for primates. Therefore, the application of this theory for neurological patients is well-grounded. Tinson and colleagues (1989) observed that stroke patients tend to spend most of their time disengaged, with inactive tasks. Thus, brain damage can lead to a reduction of activity and a withdrawal from participating in social activities. In this context, Rose and colleagues (1998) view VR as a means for producing an enriched environment, which can even overcome some practical problems of real environments.

2. VIRTUAL REALITY

In the area of cognitive rehabilitation VR systems are becoming more sought after because they provide a flexible means to address the aforementioned practical problems. Foremost, virtual reality replaces a real environment with its virtual counterpart, which eliminates the need for leaving the clinic to experience various situations. This is not only important for in-patients, it is even more important in the chronic stage for clients still in need of intensive therapy to cope with situation in daily life. Being able to remain in the wider environment of the clinic, and stressing the non-real nature of the virtual world imparts an underlying feeling of safety, even in case of being confronted with less positive albeit possible daily situations. This kind of acting without feeling threatened is a key factor in therapies (Botella, 1998).

VR also allows the design of environments which are specifically adapted to the basic requirements for a particular type of therapy. This also involves the possibility of excluding unwanted, negative aspects or effects and emphasising beneficial ones which can further strengthen the abovementioned feeling of safety.

In addition to these attributes, VR applications allow for the tackling a critical situation or problem in a graded fashion. For example, in the case of phobias, the formidable nature of the possible situations can be organised according to a graded chain of growing intensity, through which the patient can progress from the easiest tasks to the most difficult ones. Such a gradation is difficult to reproduce in an on-field exercise. In real life situations there is much less control over possible random, unwanted, negative incidents.

Role-playing is a basic exercise in many forms of therapeutic intervention and therapy settings. In such a situation VR can assist the therapist and in doing so it almost becomes another participant in the play besides the patient. This latter case is also one of the best examples, in which the attribute of flexibility can be highlighted. The crucial role of flexibility, which is a central theme of this paper, is also prevalent for both of the aforementioned main applications, namely for diagnostic and therapeutic purposes.

If we conceptualize a virtual environment with bounded surroundings, in terms of its geometry, texture, sounds, it limits the appearance generated by the software. Turning to the dynamics of this artificial world, pre-programmed events lead to the anticipation of the whole situation, i.e. the plot, which heavily influences the behaviour of the patient. Together with the former bounded appearance this could also make the whole exercise very boring for the patient and he/she would lose every motivation for participating.

The design of this type of virtual environment lacks flexibility, which is the crux of the problem. It will result in some sort of standardisation, based on statistical evaluations. It results in standard situations, environments to be displayed, but also standardised patients and human needs. All in all, it appears to match the requirements for a majority of members of a given group, for example the whole community of patients with a particular type of aphasia all over the world. But is it really the case? Or, does it match the needs of a non-existent standard human?

On the other hand, a virtual environment enhanced by flexibility enables us to even depict the users very own real life environment in addition to artificial ones designed by the therapist in user friendly way. Regarding the dynamics of such worlds, there is the possibility of not only customising plots, but also having a real-time control over the entire therapy session, i.e. throughout the therapy process. In summary, we can state that an appropriate level of flexibility allows us to refine the program according to the specific needs of each individual patient. It provides a detailed, accurate display of the actual problem, thus providing the therapist with a tool for developing adequate diagnostic environments or therapy sessions. Such a flexible tool allows for on-line follow-ups of the findings and performance during single sessions and over time in the recovery from aphasia. The whole process becomes more pleasant for the patient.

3. CONCEPT OF THE VIRTUAL CLASSROOM AND VIRTUAL THERAPY ROOM

Following the idea of the virtual classroom put forward by Rizzo (2004) and his colleagues, who measured distracting factors during class-work, we dedicated our class to the treatment of the fear of public speaking in primary schools. Stage anxiety was the topic of another software development which was investigated by Slater and colleagues (1999). They prepared a virtual conference room with a number of avatars and assessed the reaction of the user while presenting to various audiences. One audience paid no attention to the speaker throughout the whole presentation; another audience reacted in a negative manner; and a third audience changed from negative to positive at halfway through the presentation. The authors concluded that co-presence, i.e. the feeling of being with others, seemed to amplify both the negative and the positive sentiments of the situation. More importantly, a higher perceived interest from the audience reduced the anxiety experienced when speaking in public.

Our virtual environment, titled Adjustable Virtual Class (AVC) (cf. Geiszt et al., 2006) was designed on this basis. AVC is characterized by its high degree of flexibility. The situation consists of solving a task on the blackboard, while a living classroom of classmates and a teacher comments on the users' actions. This software package consisted of two essential components: an Editor to set up the classroom with regard to its spatial organisation, textures, sounds and tasks, and a Viewer to immerse the prepared environments. As opposed to Slater's abovementioned work, the avatars in our program are under real-time control of a supervising trainee, and are able to perform four speech utterances accompanied by basic gestures. A valuable potential application of the program is its use as a direct learning tool through adequately designed blackboard tasks.

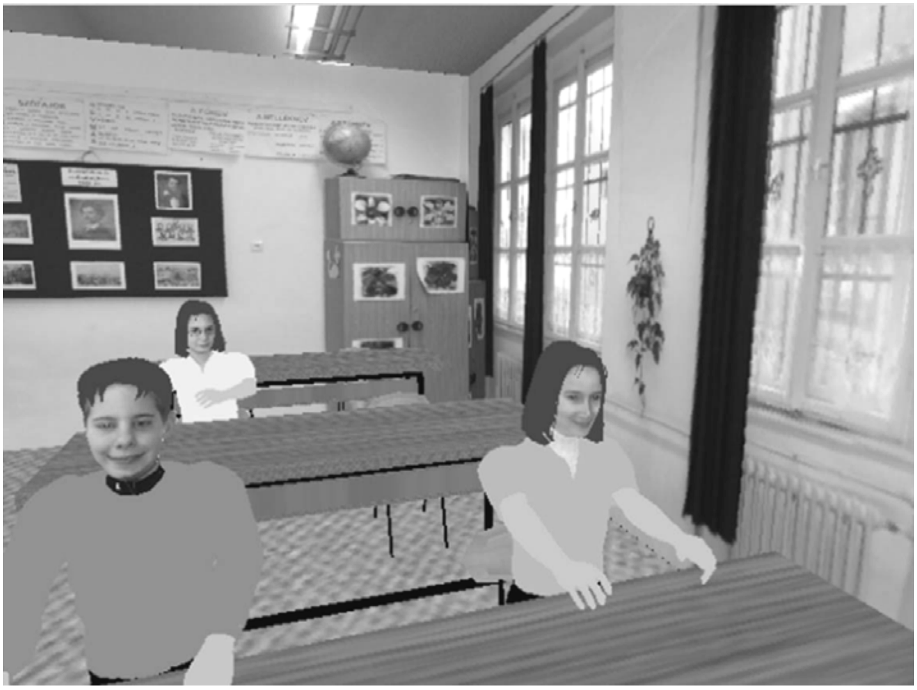


Figure 1. *Virtual students sitting in the Adjustable Virtual Classroom.*

The Virtual Therapy Room (VTR) is our second software based on the principle of flexibility. It follows the pattern laid out in the AVC. Its aim is to serve as an innovative procedure for the provision of language therapy to single aphasic clients or to several clients in a group therapy situation and for developing and testing language tasks to be performed in a VR setting. The VTR requires a different concept for the design of the environment and the situations to be displayed.

The target group for the VTR applications consists of brain-damaged persons who after having suffered a stroke present with a language impairment termed ‘aphasia’. The clinical description of the aphasic symptoms varies according to the type – Broca’s, Wernicke’s, anomic, global - and the degree of severity of the acquired language impairment.

The traditional language rehabilitation process for aphasics involves single patient and group therapy sessions designed to address each aphasic’s specific language processing difficulties. VR applications of language therapy for both single patients and group therapy are being developed in our program. In both cases, the therapist can build up an environment with the Editor part, then the patient can immerse in it with a head mounted display. During experiencing the virtual environment he/she has to answer questions appearing on the blackboard. The questions are accompanied by picture stimuli. The picture stimuli used for all of the language therapy tasks are taken from *The Everyday Life Activities (ELA) Photo Series* (Stark, 1992, 1995, 1997, 1998 and 2003). The therapist monitors the aphasic client’s doings via a display and switches between the tasks and controls and changes the actions of the present virtual persons.

Figure 2. *Initial version of the Virtual Therapy Room layout.*

The group session is similar to that of the AVC setting, but there is a more direct interaction between the single users and the virtual group members in the form of cooperation or even competition among the members. Although all the members of the virtual community are sitting around a round table, each member has his/her turn to answer a question appearing on the blackboard. If a certain amount of time passes without a solution being provided by the participant whose turn it is, other members are free to answer or to provide some cues or prompts. Thus, in our software the supervising therapist can make them perform those actions.

This competition-type motivation is not part of the single patient session, since here the focus is on situational exercises, i.e. systematic, linguistically structured tasks, for which the patient is alone with the virtual therapist. Thus, it obviously follows that the real supervisor controls only the virtual therapist and the tasks, although the interaction is more intensive between the user and the virtual therapist than in the virtual group therapy setting.

A further extension of the task is to enhance the learning process and thus strengthen the language stimulus for each task item. This is done by shifting of the viewpoint from first person view to a side view after every correct answer is given. In this shifting of viewpoint the user observes his/her own avatar giving the correct answer.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

4. DEVELOPMENT

The development of such an environment comprises the preparation of elements from which a world is built up. These elements (or building blocks) of the virtual world are selected according to the needs of the persons who will participate in the VR therapy. Then by way of observation, movement in this world has to be programmed. Lastly, in the case of the Virtual Therapy Room, the program records data for long-term evaluation of the performance of the individual users. In the following, we will concentrate on the development process of this program.

The most important part for the creation of building elements is the modelling process, which can be divided into two parts on the basis of the type of the object modelled: character modelling and the preparation of the inanimate surroundings. The first was done with Poser 4.0, preparing a number of characters, while the second was carried out with 3D Studio Max. Due to the fact that every little motion - from the standpoint of a virtual world - requires a great deal of computation in order to build up the part of the environment which is on view, we had to reduce the number of polygons used in the objects. This was done mostly manually, especially during the pre-sketch of the inanimate objects, while in the poser-generated figures, we used the older 4.0 version of the software due to the lower number of polygons used in that version instead of than in the new 6.0 version.

The second step was the refinement of texture-maps with UVMapper Classic, and for the pictures with PhotoShop 6.0.

Other pre-made elements contain recorded audio materials for avatar utterances and background noise, photos for texturing the models prepared or giving support through illustrating tasks.

The task of the frame software is not only of building up the virtual environment from the prepared models, textures, animations, audio material, and putting it into motion, but also providing a way to prepare and save custom therapy rooms, and save data including audio record of the patients' speech utterances while immersed in the virtual world. It was programmed in C++ language with the help of Irrilicht 3D Game Engine.

In spite of its name, this latter is not an engine for games but rather a so-called library: it is a collection of a number of inbuilt functions, procedures, which aid in the preparation of 3D applications. It is a platform-independent library, thus, it can work with both versions of Linux and Windows. Irrilicht 3D is freeware and can be freely modified.

Most important, from our point of view, is that Irrilicht supports a high number of 3D file types, including the standard object (.obj), Microsoft DirectX (.x), both of which were suited for our development process. The structure of the building elements comprising a prepared therapy room are saved with the help of XML in a properly structured form.

5. USE OF THE VIRTUAL THERAPY ROOM

As it was mentioned before the program has a dual structure. In the first part, i.e. Editor, there is an array of adjustable features. The options menu is divided into two sections: the General settings, comprising attributes of visualisation and room set-up, and the User menu in which the patient data-management is processed.

In the General settings section, the usual entries can be found including screen resolution, colour depth, version of DirectX, etc., and the language of the Editor interface. At present, it is able to communicate in both Hungarian and English.

Under these settings, a list of features concerning the avatars can be found. The model itself and the textures can be imported from files, as well as the audio material for a total of five speech utterances. The suggested picture format is JPEG, while for the audio it is WAW.

The Exercises section of the General settings menu is responsible for the setup of the task to be solved by the user while 'acting' in the virtual environment. Both the question and the answer can be accompanied by text and with a picture appearing simultaneously or following the written presentation on the blackboard of the virtual classroom. Audio material can be imported and associated to the task item presented on the virtual blackboard. The text of the question and the answer is limited to 200 characters.

The therapy room section is responsible for setting up the community. The floor plan itself is listed here as well as the number of the patients and therapists along with their whole representation (model with textures and voice). The entries concerning the spatial structure of the room itself are also listed here. A pre-made model can be selected for the room, in which the coordinates of the main objects must also be given.

One of these objects is the blackboard the position of which is very important, because the position of the written texts depend on it. The text includes e.g.: instructions, a task item or question and the responses to chose from (4 picture multiple choice). The position of the therapists and the patients must also be given. Particular attention is paid to their order, since this determines their identifier number. Lastly, the initial position and direction of the camera is also set up here. The therapy room was conceptualized according to an actual therapy room. It was structured this way in order to accommodate both single and group sessions. The therapist selects whether a single session or a group session will be carried out.

The User menu covers the data-management of the users' classified under self-chosen nicknames for avoiding any possibility of violation of human rights. This nickname is used for identification of a person, whose allocated set of data contains gender, type of aphasia, time post onset of aphasia and the audio records.

The task of the Viewer part is to display the virtual therapy room settings which were determined in the Editor. While the user is immersed in the virtual environment by means of a Head Mounted Display, a therapist can follow his/her actions in the virtual class through the monitor and the therapist can give instructions, comments, feedback. In addition, the therapist is able to control the virtual environment as it was mentioned before. Numerous shortcuts are offered by the software or the control of the environment.

Figure 3. *A language therapy task in the Virtual Therapy Room.*

For example the function keys from F1 to F6 select the virtual patients with the corresponding number, while F12 stands for the virtual therapist. If an avatar is selected, by pressing the numbers 1 to 5, an activity, which was set before in the Editor part, is associated with the avatar, and this activity starts immediately..

Another set of keys is related to the handling of tasks. Numeric + and – shows or hides an exercise, while the right and left cursor keys move on and back in the list of tasks and the down key selects one randomly.

Through these options the supervising therapist can build up any situation she/he wants, and can control it in real time. Due to the fact that the aphasic patients tend to require more time to respond, the supervising person has enough time to plan for the upcoming interaction and to direct the communication in a specific direction. Even dynamic group work can be built up without any additional physical members besides the supervising therapist and the single user.

These features are particularly important for the group sessions, where the aforementioned competition and support-like interactions set stronger group dynamic and social interactive factors into motion, and motivate and/or encourage the user to participate more actively in the problem-solving activity. In this way the aforementioned enriched environment can be guided and enhanced. This is a crucial factor in the case of stroke patients suffering from various types of aphasia (Broca's, Wernicke's, anomic) with varying degrees of severity of impairment (moderate, mild, minimal).

6. PILOT TEST OF THE VIRTUAL THERAPY ROOM

Following the development of the program, items for three language tasks were incorporated in the VR system: sentence comprehension, a single word processing task consisting of providing definitions to presented stimuli and a picture description task. The first two tasks consist of items of varying complexity. The number of tasks and the number of task items (n=20) will be expanded on as more tasks are developed. For this pilot study, five aphasic clients (NE, EK, WS, HW and TH) were selected for testing the preliminary version of the VR system in a single therapy setting and in the group setting. The clients were selected from aphasics who attend the 'Aphasia Club', at the Department of Linguistics and Communication Research of the Austrian Academy of Sciences. Aphasic clients with global aphasia were not included in this project, because their language impairment was too severe for them to understand the instructions to the tasks.

In single sessions each of the aphasic clients were asked to participate in this pilot study and were extensively instructed on the new method. They were acquainted with the Head Mounted Display and given adequate time to get used to this new device. After they felt comfortable wearing the Head Mounted Display the program was started. The three language tasks were carried out in a single session. Following this each client was asked for his opinion of the whole procedure. In a second session the group therapy program consisting of more interactive language tasks and addressing such pragmatic issues as turn-taking in a dynamic, group situation were performed by all five aphasics. (It must be stressed that the language tasks currently included in the VTR application are provisional and require adaptation based on the preliminary applications.)

7. SUMMARY

The two programs proved well on pilot testing: the AVC was tested in two primary schools, one with average and one with students with special needs. It is already in use in primary schools in Veszprém, Hungary. The participating teachers claim that the program is more than just a good support for their work, while the children found it to be great fun. With regard to the virtual therapy room, the initial applications of this innovative approach resulted in a high acceptance by the participants of the pilot study. The VR applications with these language impaired clients, who had received only traditional language therapy to date, revealed interesting findings regarding the language therapy process itself and also with respect to using the VR equipment. These include very different responses by the participants in the learning process in a completely different therapy situation. The results are presently being analyzed for adaptations to the applications and for future VR developments. Overall, the therapists participating in the pilot study responded to the new therapy setting positively. The fact that they could build up customized situations for group therapy sessions, ranging from friendly, cooperative ones to competition-based situations with real-time control over each virtual participant also created a learning situation for them. Interestingly, this learning aspect for both the therapist and the aphasic client relativised their 'client' versus 'therapist' role in the therapy process.

Due to its flexibility, the large number of customised elements, the virtual therapy room is an appropriate tool for both diagnosing and providing therapy for specific cognitive deficits, although the latter was the focus of our work. In summary, flexibility not only plays a crucial role when providing systematic linguistically based language therapy to aphasic clients on a one-to-one basis, it is also an important feature of VR therapy applications for aphasic clients. Future developments of the AVC and VTR - with particular emphasis on expanding the flexibility of the various components - will result in innovative procedures for providing language therapy and for diagnosing specific neuropsychological deficits in the future. Thus, returning to the title of this paper: Yes, we do need high-scale flexibility in virtual therapies.

Acknowledgements. The authors would like thank the Austrian Science and Research Liaison Office (project number: 2005.N.VII) for their support in the development of the Virtual Therapy Room.

8. REFERENCES

C Botella, C Perpina, R M Banos and A Garcia-Palacios (1998a), Virtual Reality: A New Clinical Setting Lab., In *Virtual Environments in Clinical Psychology and Neuroscienceice*, (G Riva, B Wiederhold and E Molinari, Eds), Ios Press, Amsterdam.; Available at: http://www.cybertherapy.info/pages/book2.htm (last checked on 15.07.2006)

C Botella, R Banos, C Perpina, H Villa, M Alcaniz and A Rey (1998b), Virtual Reality Treatment of Claustrophobia: A Case Report, in *Behavioural Research and Therapy*, **2**, 36, 239-246.

Z Geiszt, C Sik Lányi (2006), A Fully Adjustable Virtual Classroom, *Lecture Notes in Computer Science, LNCS 4061*, ICCHP 2006, Springer Verlag Berlin-Heidelberg, pp: 823-830.

G Kempermann, H G Kuhn and F H Gage (1997), More Hippocampal Neurons in Adult Mice Living in an Enriched Environment, in *Nature*, 386, pp 493-495.

V Laky and C Sik Lányi (2003), To develop Virtual Reality worlds for treating agoraphobia and acrophobia, in *Proc*. *Virtual Reality International Conference - VRIC*, Laval, pp. 127-133.

M North and S North (1996), Virtual Reality Psychotherapy, in *The Journal of Medicine and Virtual Reality*, *1*, 28-32.

L Pugnetti, L Mendozzi, E Barbieri and A Motta (1998), VR Experience with Neurological Patients: Basic Cost/Benefit Issues, In *Virtual Environments in Clinical Psychology and Neuroscienceice*, (G Riva, B Wiederhold and E Molinari, Eds.), Amsterdam: Ios Press; Available at: *http://www.cybertherapy.info/ves%20in%20clinical%20psycho/Pugnetti.doc.pdf* (last checked on 15.07.2006).

M J Renner and M R Rosenzweig (1987), *Enriched and Impoverished Environments: Effects on Brain and Behaviour*, Springer-Verlag, New York

A A Rizzo, Y T Bowerly, C Shahabi, J G Buckwalter, D Klimchuk, R Mitura (2004), Diagnosing Attention Disorders in a Virtual Classroom, in *IEEE Computer* 37, pp. 87-89.

F D Rose, E A Attree, B M Brooks and D A Johnson (1998), Virtual Environments in Brain Damage Rehabilitation: A Rational from Basic Neuroscience, In *Virtual Environments in Clinical Psychology and Neuroscienceice*, (G Riva, B Wiederhold and E Molinari, Eds), Ios Press, Amsterdam.

B Rothbaum, L Hodges, R Kooper, D Opdyke, M Willford and M North (1995), Effectiveness of Computer-Generated (Virtual Reality) Graded Exposure in the Treatment of Acrophobia, in *American Journal of Psychiatry*, 152, 626-628.

T Shallice, and P W Burgess (1991), Deficits in Strategy Application Following Frontal Lobe Damage in Man, in *Brain, 114*, 727-741.

C Sik Lányi, V Simon, L Simon and V Laky (2004), Using Virtual Public Transport for Treating Phobias, *Proc 5th International Conference on Disability, Virtual Reality and Associated Technologies - ICDVRAT*, Oxford, pp. 57-62.

M Slater, D-P Pertaub and A Steed (1999), Public Speaking in Virtual Reality: Facing an Audience of Avatars, *IEEE Computer Graphics and Applications* **2**, *19*, pp. 6-9.

J Stark (1992-1998), Everyday Life Activities Photo Series, Manuals to Set 1 to Set 3. Bösmüller / Jentzsch, Vienna.

J Stark (2003), Everyday Life Activities Object Photo Series Set 1, Jentzsch, Vienna.

D J Tinson (1989), How Stroke Patients Spend their Days: An Observational Study of the Treatment Regime Offered to Patients with Movement Disorders in Hospitals Following Stroke, *International Disability Studies***,** 11, 45-49.

Use of virtual reality as therapeutic tool for behavioural exposure in the ambit of social anxiety disorder treatment

H Grillon[1], F Riquier [2], B Herbelin[3] and D Thalmann[4]

[1,4]Virtual Reality Laboratory, EPFL,
IC ISIM, Station 14, 1015 Lausanne, SWITZERLAND

[2]Pychiatrist and psychotherapist,
Rue du Valentin 61, 1004 Lausanne, SWITZERLAND

[3]Computer Science Department, Aalborg University Esbjerg,
Niels Bohrs vej 8, 6705 Esbjerg, DENMARK

[1]*helena.grillon@epfl.ch*, [2]*fobia@bluewin.ch*, [3]*bh@aaue.dk*, [4]*daniel.thalmann@epfl.ch*

[1,4]*vrlab.epfl.ch,* [3]*www.aaue.dk*

ABSTRACT

We hereby present a study whose aim is to evaluate the efficiency and flexibility of virtual reality as a therapeutic tool in the confines of a social phobia behavioural therapeutic program. Our research protocol, accepted by the ethical commission of the cantonal hospices' psychiatry service, is identical in content and structure for each patient. This study's second goal is to use the confines of virtual exposure to objectively evaluate a specific parameter present in social phobia, namely eye contact avoidance, by using an eye-tracking system. Analysis of our results shows that there is a tendency to improvement in both the questionnaires and eye contact avoidance.

1. INTRODUCTION

The goal of our study is to define a therapeutic program for social anxiety disorders using virtual reality (VR) and to assess its efficiency in order to confirm that VR is a promising tool for psychotherapists in the ambit of social phobia treatment. We equally introduce our eye-tracking device as a new tool for the assessment of social phobia and present our preliminary results obtained with it.

Nowadays, exposure to VR presents itself as an alternative to standard in vivo exposures in the context of cognitive and behavioural therapy (CBT). As of today, several studies have been conducted regarding the use of VR in the treatment of social phobia (North et al., 1998; Pertaub et al., 2001; Pertaub et al., 2002; Harris et al., 2002), all leading to the conclusion that VR immersion seems adequate for such treatments but evaluated on a limited sized cohort. Anderson's study (Anderson et al., 2003) equally evaluates the treatment on a small cohort (2 people). James (James et al., 2003) concludes that a socially demanding VR environment is more anxiety provoking for a phobic than a non-socially demanding one. Slater (Slater et al., 2004) demonstrates that the difference in impact between an empty room and a room with avatars is more important in the case of phobics than in that of non phobics. In her study, Klinger (Klinger et al., 2005) concludes that both VRT (Virtual Reality Treatment) and CBT treatments are clinically valid and that the difference between the two is trivial.

Regarding eye contact, Horley (Horley et al. 2001) conducted a study on visual scanpath over 15 social phobic subjects and 15 non phobic subjects. Her results suggest that the avoidance of salient facial features is an important marker of social phobia.

As preliminary work, we have conducted a study during which we exposed subjects to a VR situation representing a 3-dimensional audience composed of emergent gazes in the dark and surrounding the subject (Herbelin et al., 2002; Riquier et al., 2002). We experimentally confirmed that the audience was able to provoke more anxiety to social phobics than to non phobics and emitted the hypothesis that eye contact is an important factor of social phobia. We therefore developed and experimented with an eye-tracking setup integrated in the VRE system (Herbelin 2005) and concluded that eye-tracking technology could "provide therapists with an objective evaluation of gaze avoidance and can give tangible feedback to the patients to estimate their progress during gaze behaviour exercises"(p.62).

In the second section of this paper, we describe our research protocol; we then present our results in the third section and conclude in section 4.

2. RESEARCH PROTOCOL

VR offers anxiety provoking scenarios which are difficult to access and are not easily available in real life. As an example, it would be extremely difficult for a therapist to fill his/her office with spiders in order to treat a patient. Equally, it would be extremely expensive and time consuming to repeatedly take a patient on an airplane in order to treat him/her against fear of flights. VR also allows repeating exposures without limitations. For example, a job interview is an accessible but exceptional situation. It would be difficult to have to do a job interview every week, as a habituation exercise.

In order to evaluate the efficiency and the potential of VRE, we use one of the social situations which are most characteristic of social phobia: the fear of public speaking according to Hofmann's model (1997). To this end, we have conceived a framework based that model. We replace the group exposure situations proposed in this therapy by individual exposure sessions to different virtual public speaking situations (Figures 1, 2, 3, 4). Phobic subjects usually recourse to avoidance strategies concerning fearful situations. The aim of these exposures is to confront the subject to his/her fear and by habituation, make him/her cope with anxiety instead of avoiding it.

Figure 1. *Job interview simulation.*

Figure 2. *Meeting in a bar.*

Figure 3. *Meeting in a cafeteria.*

Figure 4. *Speech in front of an auditory.*

In order to evaluate the efficiency of our program, we use various scales specific to social anxiety disorders at different phases of the treatment, namely the Fear Questionnaire (Marks and Matthews, 1979), the Liebowitz social anxiety questionnaire (Yao et al., 1999), the Social Interaction Self-Statement Test (Yao et al., 1998), and the Beck Depression Inventory (Beck et al., 1961). Our aim is to obtain a normalization of the score values for each subject after treatment as well as to uphold the improvement during the follow-up evaluations.

Regarding evaluation of visual contact avoidance, we use an eye-tracking system, coupled with exposure to two virtual scenes with a 3 minute verbal expression exercise in each case. We evaluate eye contact avoidance from the recording of the pupil by an eye-tracking system during exposure to the virtual scene. We seat the subject in front of a large back projection screen and equip him/her with the eye-tracking device. We then record the pupil movement at the rate of 60Hz (60 measures per second). Finally, we analyze this recording and materialize the data as a map showing the zones swept by the gaze as well as the lapse of time contact lasted.

Our eye-tracking system allows us to put into evidence and analyze in the virtual environment's zones which are looked at by the subject in real time (Figure 5). We evaluate the eye contact in a pre-therapeutic phase and at the end of the treatment. We can therefore estimate the possible effect of the treatment on this parameter.

Figure 5. *Eye-tracking system visual demonstration [Herbelin, 2005].*

This project is a clinical experiment following an A-B protocol. During the A phase − the non-intervention phase − we establish and analyse the target symptoms evolution curve through 3 evaluation sessions. During the B phase − the intervention phase − we expose the subjects to anxiety provoking situations through an HMD (Head Mounted Display) on a weekly basis, during 8 weeks; each session lasting approximately 30 minutes of which 10 in the HMD.

The study is conducted over 8 subjects recruited via a mailing sent to students in 2nd and 3rd years of college and via ambulatory consultations specialized in anxiety disorders. We admitted these subjects to participate to the study after a structured interview regarding socio-demographical variables as well as psychiatric and medical antecedents. We led a diagnostic according to the *DSM-IV*'s 5 axis for each subject and presented them with the M.I.N.I. (Mini-International Neuropsychiatric Interview) (Sheehan et al., 1998) in order to verify the prevalence of social phobia and the absence of comorbidity. Originally, they were 10 but two dropped out during the A phase of the treatment.

During the A phase, at weeks −2, −1 and 0, we asked the subjects to fill in the above mentioned questionnaires (Fear Questionnaire, Liebowitz scale, SISST and BDI). We then analysed them and averaged the results we obtained over the three weeks in order to obtain a before-treatment value for each subject and each scale.

Between phase A and phase B, subjects participated in a group session without VR. We instructed them on social phobia and asked them, one after the other, to give a speech on what they had learned about this anxiety disorder.

For the B phase, we asked the subjects to mention 8 social situations and to classify them from least to most anxiety provoking. We then exposed them to various virtual situations throughout the 8 HMD sessions, each more anxiety provoking than the previous one. The proposed virtual scenes were:

- In an office, facing one man or one woman (fig. 1)
- In an office, facing 5 people
- In an auditory, facing one man or one woman
- In an auditory, facing approximately 20 people or sitting at the back of the room (fig. 4)
- In a cafeteria, facing one person but with many people around (fig. 3)
- In a bar, facing one person but with many people around (fig. 2)

It is known that to deal with a person of the opposite sex is a typical trait of social phobia. We therefore consider the office with one person (man or woman) and the auditory with one person (man or woman) as four different situations.

We set the virtual characters in each of these scenes with a number of pre-recorded sentences which can be triggered by the psychotherapist to respectively begin, continue and end the speech session. We equally set up the virtual characters with facial animation corresponding to each of the pre-recorded sentences. We have noticed that vocal interruptions from the therapist during HMD exposure created breaks in presence. We therefore avoided these by making our virtual characters talk instead. Finally, we set up our characters with a "look at" function which allows them to make eye contact at all time and more specifically when talking to the exposed subject.

For the first HMD session, we asked each subject to present social phobia once again, as they did for the group session. Then, each week, we asked them to prepare the following week's session. As homework, they had to prepare the following week's speech in front of a mirror in order to auto-evaluate their body language. Sessions 2 to 8 consisted in the following themes:

- Session 2: talk about hobbies
- Session 3: talk about professional or educational activity
- Session 4: talk about a memorable event
- Session 5: talk about a dramatic situation
- Session 6: talk about a conflict situation
- Session 7: talk about anxiety related to love
- Session 8: we gave them documents discussing "efficient communication" and asked them to talk about these documents as if giving a lecture

We gave them these situations in this specific order for each to be more personal than the previous, and therefore, more anxiety provoking. However, since each subject isn't affected by each situation in the same way, we modulated these according to each subject. As an example, some subjects recited a poem or sang a song for session 7 because talking about their love life wasn't sufficiently anxiety provoking.

As in Hofmann's therapeutic program, we asked, as homework, for the subjects to prepare and repeat speaking exercises in front of a mirror. We also asked them to try to decrease their avoidance behaviours in real life. Finally, we also asked them to fill in a "fearful situation" document in which they exposed the anxiety provoking situations to which they have been confronted in the past week as well as the degree of avoidance, the degree of anxiety, and the satisfaction felt throughout this experience. We then used this document as base to each weekly discussion.

We asked the subjects to fill in the same 4 questionnaires as during the A phase of the treatment at week 5, half way through the treatment and once again at week 9, after the end of the HMD sessions.

We conducted the eye-tracking tests before and after the HMD sessions in order to analyze the progression in eye contact before and after treatment. Before starting with the first session, we exposed the subjects to a 5 minute habituation session. We asked them to write their name on the back projection screen with their eye. This was done in order to habituate them to the equipment and to relax before exposure by playing a game. During both sessions, we set the subjects in front of two different scenes, facing one man sitting in his office (fig. 1) and facing an auditory containing an audience of approximately 20 virtual actors (fig. 4). We did these two recordings in order to check whether the eye contact attitude is the same in different situations or not.

Our protocol equally plans another group session, which has not taken place yet. During this group session, we will ask them to present their lecture on efficient communication once again, in front of the others. This will equally allow us to see if there is a difference in attitude between HMD session and in vivo group session.

Finally, our protocol plans a follow-up session, at week 24 to verify the presence of continued evolution.

3. RESULTS

The results we present here only concern 7 people out of the 8 since the last subject has not yet been through all the HMD sessions; we therefore do not have his post-treatment results yet.

We note a general improvement for most subjects through the analysis of the various questionnaires. We can equally see that the tendencies for each subject are repeated throughout all questionnaires.

We also note that visual contact avoidance decreases. Our results show that the subjects present less avoidance behaviour after treatment than before treatment.

We equally note that one person out of the 8 does not follow the improvement pattern, on the contrary. We can see that the subject 'G''s evolution is the opposite of all other subjects'. If we took out this subject from our study, we would have a much better mean evolution.

3.1 Questionnaire Analysis

In the mean, we can see that the social phobia score for the Fear Questionnaire has decreased from 22.3 to 16.4 for women and from 19.8 to 16.5 for men. The norms to this questionnaire for phobic subjects are of respectively 15.94(8.96) to 23.4(8.4) for women and 21.4(5.44) to 24.4(8.0) for men (Bouvard, Cottraux, 2002). Our results are therefore promising.

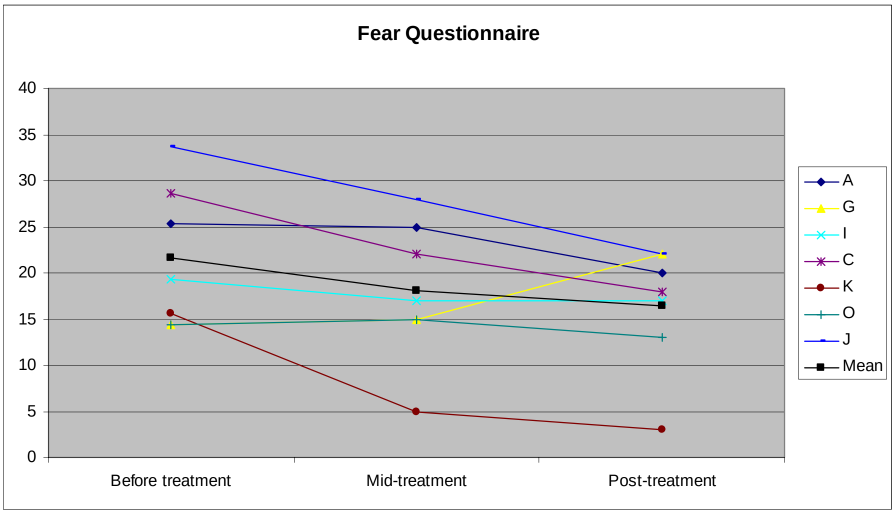


Figure 6. *Graphical results to the Fear Questionnaire.*

Regarding the Liebowitz questionnaire, the mean score decreased from 73.3 to 56.7 (Bouvard, Cottraux, 2002). Knowing that the norm for social phobic subjects is of 67.2(27.5), our results are equally promising.

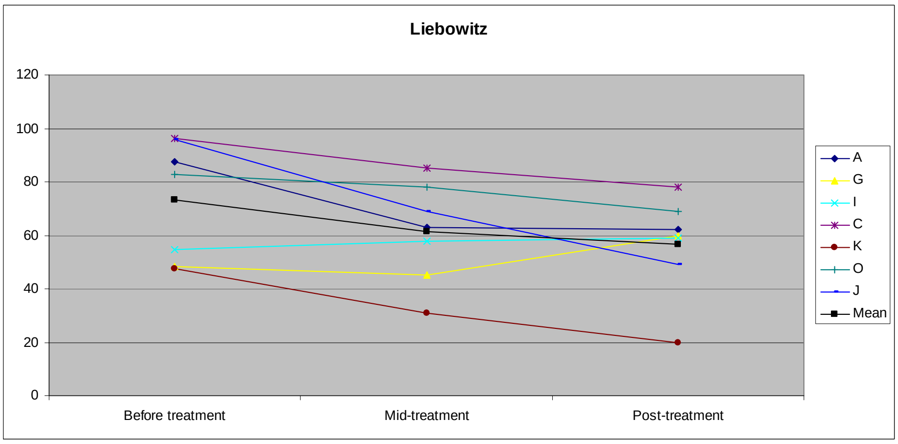


Figure 7. *Graphical results to the Liebowitz Questionnaire.*

In the mean, the score to the SISST positive thoughts has evolved from 37 to 42.6 and the score to the SISST negative thoughts, decreased from 46.9 to 38.6. For phobic subjects, the norms to this test are of 36.93(7.40) and 53.46(9.11) for positive and negative thoughts respectively (Bouvard, Cottraux, 2002). Once again, these results confirm those to the previous tests.

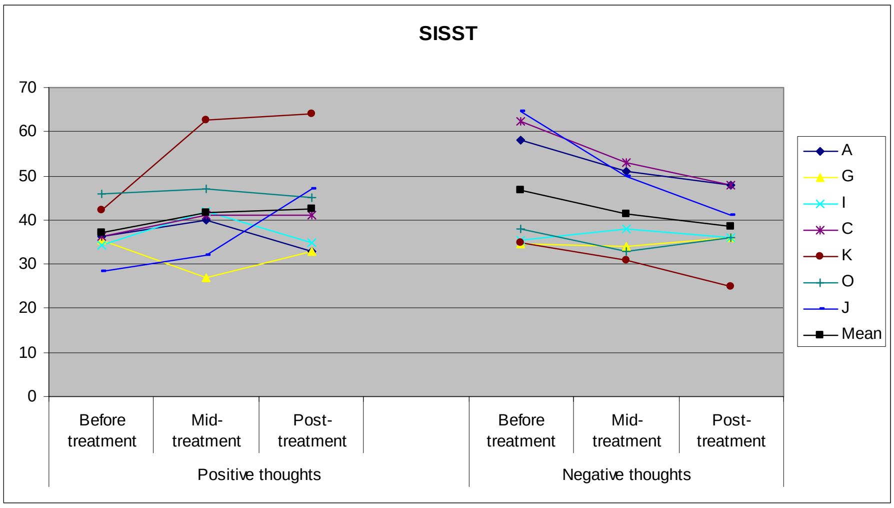


Figure 8. *Graphical results to the SISST Questionnaire.*

For the BDI questionnaire, the mean score has decreased from 9.7 to 5.3. Knowing that a score of 18.7(10.2) denotes a slight depression and that a score of 10.9(8.1) denotes no depression whatsoever (Bouvard, Cottraux, 2002), we can conclude that our subjects were not depressive should it be at the beginning of the treatment or after its end.

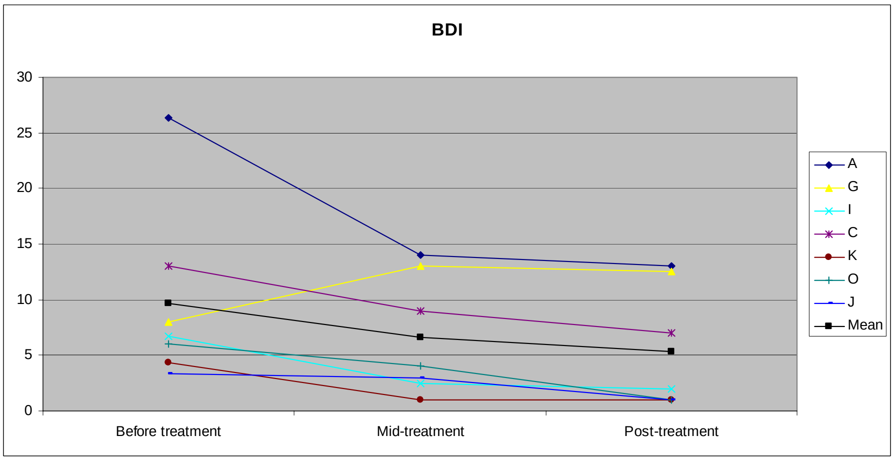


Figure 9. *Graphical results to the BDI Questionnaire.*

Figure 10. *Eye-tracking results for the interview simulation,* ***Left****: before treatment,* ***Right****: after treatment.*

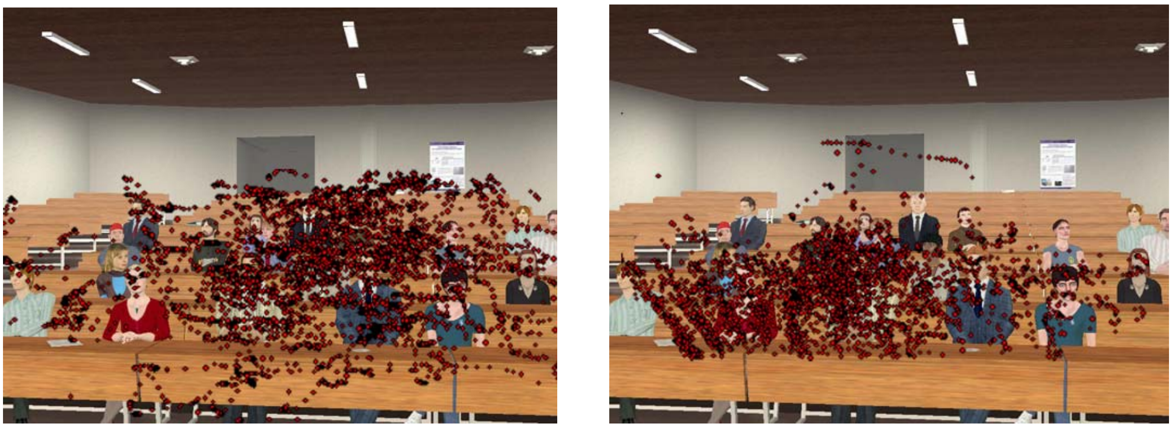


Figure 11. *Eye-tracking results for the speech simulation, **Left**: after treatment, **Right**: after treatment.*

3.2 Eye-tracking

We hereby present a case study of the outcome of eye-tracking recordings. The evolution in eye contact is different in each case but this is a representative tendency in visual contact improvement.

It is noticeable, from these recordings, that after treatment, eye contact behaviour has changed. We can notice that the virtual character's face is much more looked at after the end of the treatment than before treatment (fig. 10(a) and fig. 10(b)). We can see that the talking virtual characters (the lady in red and the man with the white sweatshirt respectively) are equally much more looked at after treatment than before treatment in the second situation (fig. 11(b) and fig. 11(b)). These results tend to show that eye contact avoidance has diminished after treatment and that looking at the talking person's face has become less difficult.

4. CONCLUSIONS AND DISCUSSION

In this paper, we have presented a study whose aim was to evaluate the efficiency of a therapeutic program using VR as a tool to treat social phobia. We have presented the VR-based protocol which we have used to conduct a study over 8 social phobic subjects. We have equally presented the results we have obtained through questionnaire analysis and eye-tracking recordings.

Firstly, we have noticed a general improvement for most subjects after having analysed all questionnaires. We have equally noticed that the tendencies for each subject were the same over the questionnaires. This leads us to think that certain people are more reactive to VRET than others. We have seen, in our results, that one of the subjects apparently wasn't affected by the treatment at all. However, due to the limited size of our sample, we cannot conclude to the validation of VR as a treatment for social phobia. Nevertheless, our results show that it could be a promising therapeutic tool and that VR demands more extensive experimentation. In order to validate this tool, another study should be conducted over a minimum of 30 subjects and with a control group. However, 50 subjects should be recruited since some of them will certainly drop out before the end of the study, as we have seen in ours.

Secondly, and according to our former experience in case studies (Herbelin 2005), we have observed a deviation of the gaze during VRE and have obtained a qualitative appreciation of the patients' gaze avoidance. After having analysed the eye-tracking recordings, we have equally noticed an improvement in eye contact avoidance. We intend to conduct more extensive research on eye contact avoidance and analyse our results in more detail, in order to correlate the results between scenes for a same subject as well as between subjects themselves. We also intend to investigate the eye scan and its speed.

By using subjective appreciation of social anxiety throughout the whole experiment (the 4 above mentioned questionnaires) as well as video recordings of the exposure sessions and the results provided by our eye-tracking system, we have noticed an improvement regarding avoidance behaviour and a decrease of anxiety with time and exposure. The follow-up with the subjects will give us the opportunity to validate our actual results regarding VR therapy on a long-term basis and its ability to continue providing results once the therapy is over.

Acknowledgements: We would like to thank Diana Felli and Caroline Nicod for having helped us conduct the HMD sessions, Mireille Clavien for the design and Benoît Le Callennec for careful proofreading. Part of the software has been developed thanks to a Swiss National Research Foundation grant.

5. REFERENCES

American Psychiatric Association (1994), *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. 4th Edition, Washington DC.

Anderson, P., Rothbaum, B.O., Hodges, L.F. (2003) *Virtual reality in the treatment of social anxiety: Two case reports.* Cognitive and Behavioral Practice 10, 240-247.

Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961) *An inventory for measuring depression.* Archives of General Psychiatry 4, 561-571.

Bouvard M., Cottraux J., (2002) *Protocoles et échelles d'évaluation en psychiatrie et en psychologie*. Masson, Paris, 3è édition.

Harris, S.R., Kemmerling, R.L., North, M.M. (2002) *Brief virtual reality therapy for public speaking anxiety.* CyberPsychology & Behavior 5(6), 543-550.

Herbelin B. (2005) *Virtual reality exposure therapy for social phobia.* Thesis N° 3351, VRLab, Faculté informatique et communications, Institut des systèmes informatiques et mutlimédias, EPFL, Switzerland.

Herbelin, B., Riquier, F., Vexo, F. and Thalmann, D. (2002) *Virtual reality in cognitive behavioral therapy: a preliminary study on social anxiety disorder.* 8th International Conference on Virtual Systems and Multimedia (VSMM 2002), Gyeongju, 25-27 September 2002

Hofmann S.G. (1997) *Self-focused exposure therapy for social phobia.* unpublished treatment manual, Boston University.

Horley K., Williams, L.M., Gonsalvez C., Gordon E. (2001) *Social phobics do not see eye to eye: A visual scanpath study of emotional expression processing.* Journal of Anxiety Disorders 17, 33-44.

James, L.K., Lin, C.Y., Steed, A., Swapp, D., Slater, M. (2003) *Social anxiety in virtual environments: results of a pilot study.* CyberPsychology & Behavior 6(3), 237-243.

Klinger, E., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., Nugues, P. (2005) *Virtual reality therapy versus cognitive behavior therapy for social phobia : A preliminary controlled study.* Cyberpsychology & Behavior 8(1), 76-88.

Marks, I.M. & Matthews, A.M. (1979) *Brief standard self-rating for phobic patients.* Behavior Research and Therapy 17, 263-267.

North, M., North, S and Coble, J.R. (1998) *Virtual reality therapy: an effective treatment for the fear of public speaking*. International Journal of Virtual Reality 3(2), 2-6.

Pertaub, D., Slater, M. and Barker, C. (2001) *An Experiment on fear of public speaking in virtual reality.* Studies in Health Technology & Informatics 81, 372-378.

Pertaub, D., Slater, M. and Barker, C. (2002) *An Experiment on fear of public speaking anxiety in response to three different types of virtual audience.* Presence: Teleoperators and Virtual Environments 11(1), 68-78.

Riquier, F., Stankovic, M. and Chevalley A.F. (2002) *Virtual gazes for social exposure: Margot and Snow White.* First International Workshop on Virtual Reality Rehabilitation (VRMHR 2002), Lausanne, 7-8 November 2002, 35-49.

Sheehan, DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R and Dunbar GC (1998) *The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10.* Journal of Clinical Psychiatry 59, 22-33.

Slater, M., Pertaub, D.P., Barker, C., Clark, D. (2004) *An Experimental Study on Fear of Public Speaking Using a Virtual Environment.* 3rd International Workshop on Virtual Rehabilitation, Lausanne, 16-17 September 2004.

Yao, S.N., Note, I., Fanget, F., Albuisson, E., Bouvard, M., Jalenques, I. and Cottraux, J. (1999) *L'anxiété sociale chez les phobiques sociaux : validation de l'échelle d'anxiété sociale de Liebowitz.* L'Encéphale 25, 429-435

Yao, S.N., Cottraux, J., Mollard, E., Albuisson, E., Note, I., Jalenques, I., Fanget, F., Robbe-Grillet, P., Dechassey, M., Ektmedjian, N., Bouvard, M. and Coudert, A.J. (1998) *The french version of the social interaction self-statement test (SISST): a validation and sensitivity study in social phobics.* Behavioral and Cognitive Psychotherapy 26, 247-259.

User-centered design driven development of a virtual reality therapy application for Iraq war combat-related post traumatic stress disorder

A A Rizzo[1], K Graap[2], J Pair[1], G Reger[3], A Treskunov[1] and T Parsons[1]

[1]Institute for Creative Technologies, University of Southern California,
13274 Fiji Way, Marina del Rey, CA. 90292, USA

[2]Virtually Better, Inc., 2450 Lawrenceville Highway, Suite 101
Decatur, Georgia 30033

[3]U.S. Army, Fort Lewis, Tacoma Washington

arizzo@usc.edu, graap@virtuallybetter.com, pair@ict.usc.edu, tparsons@ict.usc.edu

ABSTRACT

Post Traumatic Stress Disorder (PTSD) is reported to be caused by traumatic events that are outside the range of usual human experience including (but not limited to) military combat, violent personal assault, being kidnapped or taken hostage and terrorist attacks. Initial data suggests that at least 1 out of 6 Iraq War veterans are exhibiting symptoms of depression, anxiety and PTSD. Virtual Reality (VR) delivered exposure therapy for PTSD has been used with reports of positive outcomes. The aim of the current paper is to present the rationale, technical specifications, application features and user-centered design process for the development of a Virtual Iraq PTSD VR therapy application. The VR treatment environment is being created via the recycling of virtual graphic assets that were initially built for the U.S. Army-funded combat tactical simulation scenario and commercially successful X-Box game, *Full Spectrum Warrior,* in addition to other available and newly created assets. Thus far we have created a series of customizable virtual scenarios designed to represent relevant contexts for exposure therapy to be conducted in VR, including a city and desert road convoy environment. User-centered design feedback needed to iteratively evolve the system was gathered from returning Iraq War veterans in the USA and from a system in Iraq tested by an Army Combat Stress Control Team. Clinical trials are currently underway at Camp Pendleton and at the San Diego Naval Medical Center. Other sites are preparing to use the application for a variety of PTSD and VR research purposes.

1. INTRODUCTION

In 1997, Virtually Better, Inc. released the first version of the Virtual Vietnam VR scenario for use as an exposure therapy tool for treating Post Traumatic Stress Disorder (PTSD) in Vietnam veterans. This occurred over 20 years following the end of the Vietnam War. During those intervening 20 years, in spite of valiant efforts to develop and apply traditional psychotherapeutic approaches to PTSD, the progression of the disorder in some veterans significantly impacted their psychological well being, functional abilities and quality of life, as well as that of their family members and friends. The tragic nature of this disorder also had significant ramifications for the U.S. Dept. of Veteran Affairs healthcare system often leading to designations of lifelong service connected disability status. The Virtual Vietnam scenario landmarked the first time that VR was applied to the treatment of PTSD and this initial effort produced encouraging results. In the early 21st century the conflicts in Iraq and Afghanistan again drew US military personnel into combat. Hoge et al., (2004), in the first systematic study of mental health problems due to these conflicts reported that *"…The percentage of study subjects whose responses met the screening criteria for major depression, generalized anxiety, or PTSD was significantly higher after duty in Iraq (15.6 to 17.1 percent) than after duty in Afghanistan (11.2 percent) or before deployment to Iraq (9.3 percent)"* (p.13). With this history in mind, the University of Southern California Institute for Creative Technologies (ICT) initiated a project to create an immersive virtual environment system for the exposure therapy treatment of Iraq and Afghanistan War veterans diagnosed with combat-related PTSD. This PTSD VR therapy environment is being created via the recycling of graphic assets that were initially built for the U.S. Army-funded combat tactical simulation scenario and commercially successful X-Box game, *Full Spectrum Warrior*. As well, other existing and newly created assets available to ICT are being integrated into this rapidly evolving application. The presence of ICT expertise in designing combat simulations and an interdisciplinary collaboration with leading experts in the field of combat related PTSD has led to the opportunity to apply VR for this relevant clinical challenge, albeit within a tighter timeframe than the technology allowed for Vietnam era veterans with PTSD.

According to the DSM-IV (1994), PTSD is caused by traumatic events that are outside the range of usual human experiences such as military combat, violent personal assault, being kidnapped or taken hostage, rape, terrorist attack, torture, incarceration as a prisoner of war, natural or man-made disasters, automobile accidents, or being diagnosed with a life-threatening illness. The disorder also appears to be more severe and longer lasting when the event is caused by human means and design (bombings, shootings, combat, etc.). Such incidents would be distressing to almost anyone, and is usually experienced with intense fear, terror, and helplessness. Typically, the initiating event involves actual or threatened death or serious injury, or other threat to one's physical integrity; or the witnessing or awareness of an event that involves death, injury, or a threat to the physical integrity of another person. The essential feature of PTSD is the development of characteristic symptoms that may include: intrusive thoughts and flashbacks, avoidance of reminders of the traumatic event, emotional numbing, hyper-alertness, anger, isolation, anxiety, depression, substance abuse, survivor guilt, suicidal feelings and thoughts, negative self-image, memory impairment, problems with intimate relationships, emotional distance from family and others and denial of social problems. Symptoms of PTSD are often intensified when the person is exposed to stimulus cues that resemble or symbolize the original trauma in a *non-therapeutic* setting. Such *uncontrolled* cue exposure may lead the person to react with a survival mentality and mode of response that could put the patient and others at considerable risk.

Prior to the availability of VR therapy applications, the existing standard of care for PTSD was *imaginal* exposure therapy. Such treatment typically involves the graded and repeated imaginal reliving and narrative recounting of the traumatic event within the therapeutic setting. This approach is believed to provide a low-threat context where the patient can begin to therapeutically process the emotions that are relevant to the traumatic event as well as de-condition the learning cycle of the disorder via a habituation/extinction process. While the efficacy of imaginal exposure has been established in multiple studies with diverse trauma populations (Rothbaum et al., 2000; 2002), many patients are unwilling or unable to effectively visualize the traumatic event. In fact, *avoidance* of reminders of the trauma is one of the cardinal symptoms of PTSD. It is often reported that, *"...some patients refuse to engage in the treatment, and others, though they express willingness, are unable to engage their emotions or senses*." (Difede & Hoffman, 2002). Research on this aspect of PTSD treatment suggests that the inability to emotionally engage (*in imagination*) is a predictor for negative treatment outcomes (Jaycox et al., 1998). In contrast to imaginal therapy, VR is seen to provide a controlled stimulus environment that can more effectively deliver the controlled exposure to trauma cues needed for therapeutic gain. Such use and value of VR for the treatment of cognitive, emotional, psychological and motor disorders has been well specified (Glantz, Rizzo & Graap, 2003; Rizzo et al., 2004). The first use of VR for a Vietnam veteran with PTSD was reported in a case study of a 50-year-old, Caucasian male veteran meeting DSM-IV criteria for PTSD (Rothbaum et al., 1999). Results indicated post-treatment improvement on all measures of PTSD and maintenance of these gains at a 6-month follow-up. This case study was followed by an open clinical trial of VR for Vietnam veterans (Rothbaum et al., 2001). In this study, 16 male PTSD patients were exposed to two HMD-delivered virtual environments, a virtual clearing surrounded by jungle scenery and a virtual Huey helicopter, in which the therapist controlled various visual and auditory effects (e.g. rockets, explosions, day/night, yelling). After an average of 13 exposure therapy sessions over 5-7 weeks, there was a significant reduction in PTSD and related symptoms.

The early positive findings from the Vietnam studies led other groups to propose VR environments to facilitate PTSD treatment in both civilians and military personnel. For example, subsequent to the September 11, 2001 terrorist attacks on the World Trade Center in New York, Difede and Hoffman (2002) constructed a scenario in which civilians, firefighters and police officers with PTSD could be exposed in VR to relevant trauma-related events. In their first report, a case study was presented using VR to provide exposure to the trauma memory with a patient who had failed to improve with traditional imaginal exposure therapy. The authors reported significant reduction of PTSD symptoms after repeatedly exposing the patient to explosions, sound effects, virtual people jumping from the burning buildings, towers collapsing, and dust clouds and attributed this success partly due to the increased realism of the VR images as compared to the mental images the patient could generate in imagination. Positive treatment outcomes from a wait-list controlled VR study with patients who were not successful in previous imaginal therapy are currently under review by this group (Joanne Difede, personal communication, June 17, 2006). These initial positive findings have also encouraged others to begin development and initial user-centered testing of VR scenarios to treat

PTSD in survivors of war and terrorist attacks. In Portugal, where there is an estimated 25,000 survivors with PTSD from the 1961-1974 colonial wars in Mozambique, Angola and Guiné, Gamito et al. (2005) has constructed a VR "ambush" scenario using the Halo 2 game engine as a starting point. They report having recently conducted an initial user-centered test with one PTSD patient who has provided feedback suggesting the need for the construction of a system that provides more graduated delivery of anxiety provoking trigger stimuli. In Israel, Josman et al. (2005) is currently implementing a terrorist "bus bombing" PTSD treatment scenario in which the patient is positioned in an urban cafe across the street from the site where a civilian bus may explode. The system controls allow the patient to sit in the outdoor cafe and be exposed to a range of progressive conditions − from the street being empty with no bus or sound effects − to the bus passing in an uneventful manner with or without sound − to the bus arriving and exploding with full sound effects. This research has only recently commenced and no clinical efficacy data are currently available. However, based on the data that is available across the field since the advent of Virtual Vietnam, the initial results indicate that VR exposure therapy is a vital component to be included within a comprehensive treatment approach for persons with combat-related PTSD.

2. TECHNICAL BACKGROUND AND DEVELOPMENT HISTORY

In the pragmatic design of this application, minimizing costs was achieved in two ways. Whenever possible, assets from previous applications were "recycled" for use in the current project. This can be most poignantly seen in the use of Full Spectrum Warrior X-Box game art assets as a start point for developing the initial prototypes. However, as user-centered feedback guided iterative design cycles, it became apparent that we also needed to create specialized art for the scenario development. As well, an attempt was made whenever possible to use low cost commodity off the shelf equipment in order to maximize the access and availability of the finished system. The current application is designed to run on two Pentium 4 notebook computers each with 1 GB RAM, and 128 MB DirectX 9-compatible NVIDIA 3D graphics cards. The two computers are linked using a null ethernet cable. One notebook runs the therapist's control application while the second notebook drives the user's head mounted display (HMD). We evaluated HMDs from three different manufacturers for use in this application with the aim of finding an affordable display with acceptable resolution and field of view (5DT, Cyvisor, eMagin). The HMD that was chosen was the eMagin z800, which contains OLED displays capable of 800x600 (SVGA) resolution with a 40 degree diagonal field of view (http://www.emagin.com/). The major selling point for using this HMD was the presence of a built-in 3DOF head tracker. At under $600 per unit, this integrated display/tracking solution was viewed as the best option to minimize costs and maximize the access to this system by those in need. It should also be noted that while we believe that a HMD will provide the optimal level of immersion and interaction for this application, the system is fully configurable to be delivered on a standard PC monitor or within a large screen projection display format. The application is built on ICT's FlatWorld Simulation Control Architecture (FSCA) (Treskunov, Pair & Swartout, 2004). The FSCA enables a network-centric system of client displays driven by a single controller application. The controller application broadcasts user-triggered or scripted event data to the display client. The controller can introduce scripted elements to the VR environment at runtime. These scripted elements allow the controller to provide real-time client-side interaction despite potential network delays. FCSA scripting is based on the Lua programming language and provides facilities for real-time triggered events, animation and path planning. The client's real-time 3D scenes are presented using Numerical Design Limited's (NDL) Gamebryo rendering library. Pre-existing art assets were integrated using Alias' Maya 6 and Autodesk 3D Studio Max 7 with new art created primarily in Maya.

We have also added olfactory and tactile stimuli to the experience of the environment. Recently, Virtually Better, Inc. in collaboration with Envirodine, Inc. introduced the *Scent Palette* for use in conjunction with VR environments. The *Scent Palette* is a USB device that uses up to 8 smell cartridges, a series of fans, and a small air compressor to deliver scents to participants. The scents can be activated by placing location-aware triggers into the VR application. For example, a user walks by a fire and smells smoke. The smells can also be controlled by operator key presses or mouseclicks. Scents may be employed as direct stimuli (e.g., scent of burning rubber) or as cues to help immerse users in the world (e.g., ethnic food cooking). The amount of scent to be released is specified in seconds. The scents selected for this application include burning rubber, cordite, garbage, body odor, smoke, diesel fuel, Iraqi food spices, and gunpowder. Vibration is also used as an additional user sensory input. Vibration is be generated through the use of a Logitech force-feedback game control pad and through audio-tactile sound transducers from Aura Sound Inc. located beneath the patient's floor platform and seat. Audio files are customized to provide vibration consistent with relevant visual and audio stimuli in the scenario. For example, explosions are accompanied by a shaking floor. In the HUMVEE scenarios, the user's seat vibrates as the vehicle moves across uneven ground.

3. CLINICAL APPLICATION CONTROL OPTIONS: SCENARIO SETTINGS, USER PERSPECTIVES, TRIGGER STIMULI & THE CLINICAL INTERFACE

3.1 Early Prototype and Evolving Vision

Prior to acquiring the Office of Naval Research (ONR) funding required to create a comprehensive VR application to address a wide range of possible combat-related PTSD experiences, we created a prototype virtual environment designed to resemble a small middle-eastern city (see Figures 1-2). This virtual environment was designed as a proof of concept demonstrator and as a tool for initial user testing to gather feedback from both Iraq War military personnel and clinical professionals. The resulting feedback has been used to refine the city scenario and to drive development of other relevant virtual contexts. Current ONR funding has now allowed us to evolve this existing prototype into a various versions that have undergone user-centered design feedback trials at sites throughout the USA with non-PTSD soldiers who have returned from an Iraq tour of duty. User centered feedback was also collected within an Army Combat Stress Control Team in Iraq (see Figure 3). The vision for the project includes both the design of a series of diverse *scenario settings* (i.e. city, outlying village and desert convoy scenes), and the creation of methods for providing users with different *user perspective options*. These choice options when combined with real time clinician input via a "Wizard of Oz" *clinical interface* is envisioned to allow for the creation of a user experience that is specifically customized to the varied needs of patients who participate in treatment. This is an essential component for giving the therapist the capacity to modulate patient anxiety as is required for an exposure therapy approach. Such options for user experience customization and real time stimulus delivery flexibility are key elements for these types of VR exposure applications.

3.2 Scenario Settings

The software has been designed such that patients can be teleported to specific scenario settings based on a determination as to which environments most closely match the patient's needs, relevant to their individual combat related experiences. All scenario settings are adjustable for time of day or night, weather conditions and lighting illumination. The following are the scenario settings that are being created for the application:

City Scenarios – In this setting, we are creating two variations. The first city setting has the appearance of a desolate set of low populated streets comprising of old buildings, ramshackle apartments, warehouses, a mosque, factories and junkyards (see Figures 1-2). The second city setting has similar street characteristics and buildings, but is more highly populated and has more traffic activity, a marketplace, monuments and alleys with insurgents (see Figures 3-4, 16, 17).

Small Rural Village – This setting consists of a more spread out rural area containing ramshackle structures, a village center and much decay in the form of garbage, junk and wrecked or battle-damaged vehicles. It will also contain more vegetation and have a view of a desert landscape in the distance that is visible as the user passes by gaps between structures near the periphery of the village.

Desert Road – This scenario consists of paved and dirt roadways that will eventually connect the City and Village scenarios. The view from the road currently consists of desert scenery and sand dunes (see Figures 5-6) with occasional areas of vegetation, ramshackle structures, battle wreckage, debris and virtual human figures (see Figures 7-8). The scenario supports the perception of travel within a convoy or as a lone vehicle.

Building Interiors – Some of the City and Village Scenario buildings have interiors modelled that allow the user to navigate through them. These interiors will have the option of being vacant (see Figure 9) or inhabited by various numbers and types of virtual human characters.

Checkpoints – This area of the City Scenario is being constructed to resemble a traffic checkpoint with a variety of moving vehicles arriving, stopping and then moving onward into the city.

3.3 User Perspectives

The VR system is designed such that once the scenario setting is selected, it will be possible to select from a variety of user perspective and navigation options. These are being designed in order to again provide flexibility in how the interaction in the scenario settings can be customized to suit the patient's needs. User perspective options in the final system will include:

1. User walking alone on patrol from a first person perspective (see Figures 1–4, 9).
2. User walking with one soldier companion on patrol. The accompanying soldier will be animated with

a “flocking” algorithm that will place him always within a 5-meter radius of the patient and will adjust position based on collision detection with objects and structures to support a perception of realistic movement. (see Figure 10).

3. User walking with a patrol consisting of a number of companion soldiers using a similar “flocking” approach as in #2 above (see Figure 11). These flocking options are under development and will be integrated during year two of this project.

4. An adjustable user view from the perspective of being either inside of the cab of a HUMVEE or other moving vehicles or from a more exposed position in a gun turret above the roof of the vehicle (see Figures 5, 7, 8, 12). Options are provided for automated travel as a passenger through the various setting scenarios or at the driving column that allow for user control of the vehicle via the gamepad controls. The interior view will also have options for other occupant passengers that will have ambient movement.

5. User view from the perspective of being in a helicopter hovering above or moving over any of the scenario settings (see Figure 13). Night vision is also an option (see Figure 14).

In each of these user perspective options, we are considering the wisdom of having the patient possess a weapon. This will necessitate decisions as to whether the weapon will be usable to return fire when it is determined by the clinician that this would be a relevant component for the therapeutic process. Those decisions will be made based on the initial user and clinician feedback from the clinical test application.

3.4 Trigger Stimuli

The specification, creation and addition of trigger stimuli will likely be an evolving process throughout the life of the application based on relevant patient and clinician feedback. We began this part of the design process by including options that have been reported to be relevant by returning soldiers and military subject matter experts. For example, in the Hoge et al., (2004), study of self-reported anxiety, depression and PTSD-related symptomatology in Iraq War veterans, they present a listing of combat-related events that were commonly experienced in their sample. These events provided a useful starting point for conceptualizing how relevant trigger stimuli could be presented in a VR environment. Such commonly reported events included: “*Being attacked or ambushed, Receiving incoming artillery, rocket, or mortar fire, Being shot at or receiving small-arms fire, Shooting or directing fire at the enemy, Being responsible for the death of an enemy combatant, Being responsible for the death of a noncombatant, Seeing dead bodies or human remains, Handling or uncovering human remains, Seeing dead or seriously injured Americans, Knowing someone seriously injured or killed, Participating in de-mining operations, Seeing ill or injured women or children whom you were unable to help, Being wounded or injured, Had a close call, was shot or hit, but protective gear saved you, Had a buddy shot or hit who was near you, Clearing or searching homes or buildings, Engaging in hand-to-hand combat, Saved the life of a soldier or civilian.”* (p. 18). From this and other sources, we have begun our initial effort to conceptualize what is both functionally relevant and pragmatically possible to include as trigger stimuli in the virtual environment. Thus far, we have created a variety of auditory trigger stimuli (i.e., incoming mortars, weapons fire, voices, wind, etc.) that can be actuated by the clinician via mouse clicks on the clinical interface. We can also similarly trigger dynamic audiovisual events such as helicopter flyovers above the user’s position and verbal orders from a commanding officer who is gesturing in an excited manner. The creation of more complex events that can be intuitively delivered from a clinicians’ interface while providing a patient with options to interact or respond in a meaningful manner is one of the ongoing focuses in this project. Perhaps it may be of value to actually immerse the user in varying degrees of combat in which they may see members of their patrol (or themselves) get wounded or in fact have the capability to fire a weapon back at enemy combatants. However, such trigger options will require not only interface design expertise, but also clinical wisdom as to how much and what type of exposure is needed to produce a positive clinical effect. These issues will be keenly attended to in our initial clinical trials.

3.5 The Clinical Interface

In order to deliver and control all of the above features in the system, a “wizard of oz” type clinical interface was created. This interface is a key element in the application, as it needs to provide a clinician with a usable tool for selecting and placing the patient in VR scenario locations that resemble the contexts that are clinically relevant for a graduated exposure approach. As important, the clinical interface must also allow the clinician to further customize the therapy experience to the patient’s individual needs via the systematic real-

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

time delivery and control of “trigger” stimuli in the environment. This is essential for fostering the anxiety modulation needed for therapeutic habituation.

In our initial configuration, the clinician uses a separate computer monitor/mouse or tablet laptop to display and actuate the clinical interface controls. The results from our initial user feedback trials guided the interface design whereby our setup provides four quadrants in which the clinician can monitor ongoing user status information, while simultaneously directing trigger stimulus delivery (see Figure 15) in an effort to modulate appropriate levels of anxiety as required by the theory and methodology of exposure-based therapy. The overall design of the system is such that once the scenario setting is selected, the clinician can then adjust the time of day, weather options, ambient sounds, scent and vibration configurations and user perspective. Once these options are selected, the patient can experience this customized environment setting while the clinician focuses on the judicious delivery of trigger stimuli. These interface options have been designed with the aid of feedback from clinicians with the goal to provide a usable and flexible control system for conducting thoughtfully administered exposure therapy that can be readily customized to suit the needs of the patient. Although cost factors limit the creation of custom VEs specific to the unique experiences of every person, it is possible to construct flexible archetypic VR worlds for PTSD patients that lend themselves to abstraction and some degree of commonality.

4. USER CENTERED EVALUATION

User-Centered tests with the application were conducted at the Naval Medical Center–San Diego, Weill Cornell Medical College-NYC and within an Army Combat Stress Control Team in Iraq (see Figures 18-20). This feedback with non-diagnosed personnel provided feedback on the content and usability of our application to feed an iterative design process. A clinical trial version of the application built from this process is currently being tested with PTSD-diagnosed personnel at various USA sites and any available data from these clinical tests will be presented at IEEE VR2006. We have also conducted brief exposure with two volunteer PTSD patients using an early prototype. One of these patients has reported a reduction in nightmares and the other has reported that the scenario has helped him to cognitively reframe his experience with positive results. These anecdotal and informal self-report commentaries are encouraging for our continued efforts in this clinical direction.

5. CONCLUSION

War is perhaps one of the most challenging situations that a human being can experience. The physical, emotional, cognitive and psychological demands of a combat environment place enormous stress on even the best-prepared military personnel. Such stressful experiences that commonly occur in warfighting environments have a considerable likelihood for producing significant numbers of returning soldiers at risk for developing PTSD. The initial data coming from both survey studies and anecdotal observations suggest that a large number of returning soldiers from the Iraq/Afghanistan conflicts are in fact reporting symptoms that are congruent with the diagnosis of PTSD. It is our view that this situation requires our best efforts to find ways to maximize treatment access and efficacy and VR is a logical and attractive medium to use to address these aims.

One of the more foreboding findings in the Hoge et al., (2004) report, was the observation that among Iraq/Afghanistan War veterans, *“…those whose responses were positive for a mental disorder, only 23 to 40 percent sought mental health care. Those whose responses were positive for a mental disorder were twice as likely as those whose responses were negative to report concern about possible stigmatization and other barriers to seeking mental health care*.” (p. 13). While military training methodology has better prepared soldiers for combat in recent years, such hesitancy to seek treatment for difficulties that emerge upon return from combat, especially by those who may need it most, suggests an area of military mental healthcare that is in need of attention. To address this concern, perhaps a VR system for PTSD treatment could serve as a component within a reconceptualized approach to how treatment is accessed by veterans returning from combat. One option would be to integrate VR-delivered combat exposure as part of a comprehensive “assessment” program administered upon return from a tour of duty. Since past research is suggestive of differential patterns of physiological reactivity in soldiers with PTSD when exposed to combat-related stimuli (Laor et al., 1998; Keane et al., 1998), an initial procedure that integrates a VR PTSD application with psychophysiological monitoring could be of value. If indicators of such physiological reactivity are present during an initial VR exposure, a referral for continued assessment and/or care could be negotiated and/or suggested. This could be provided in a format whereby the perceived stigma of independently seeking

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

treatment could be lessened as the soldier would be simply involved in some form of "non-combat reintegration training" in a similar fashion to other designated duties to which they would participate.

VR PTSD therapy may also offer an additional attraction and promote treatment seeking by certain demographic groups in need of care. The current generation of young military personnel, having grown up with digital gaming technology, may actually be more attracted to and comfortable with participation in a VR application approach as an alternative to what is viewed as traditional "talk therapy" (even though such talk therapy would obviously occur in the course of a recommended multi-component approach for this disorder). The potential for a reduction in the perceived stigma surrounding treatment has been anecdotally reported by practitioners who use VR to treat civilians with aerophobia (Wiederhold & Wiederhold, 2004). These authors indicate that some patients have reported that prior to treatment, they had "just lived with the problem" and never considered seeking professional treatment. Upon hearing of VR therapy for fear of flying, often via popular media reports, they then sought out VR exposure treatment, typically with resulting positive outcomes.

In addition to the ethical factors that make an unequivocal case for the importance of exploring new options for assessment and treatment of combat-related PTSD, economic drivers for the healthcare system also provide incentives for investigating novel approaches in this area. As of Sept. 2004, there were 13,524 Gulf War Veterans receiving compensation for PTSD from the Dept. of Veterans Affairs (VA Fact Sheet, 2004). In addition to the direct costs for benefit compensation, medical care usage by persons with PTSD is estimated to be 60% higher than average (Marshall et al., 2000) and lost income-based tax revenues raise the "hidden" costs even higher. These figures make the initial development and continuing infrastructure costs for running PC-based VR systems pale by comparison. The military could also benefit economically by way of reduced turnover of soldiers with mild PTSD who might be more likely to reenlist if their mental health needs were addressed soon after combat in a progressive manner via earlier VR assessment and treatment. As well, such a VR tool initially developed for exposure therapy purposes, offers the potential to be "recycled" for use both in the areas of combat readiness assessment and for stress inoculation. Both of these approaches could provide measures of who might be better prepared for the emotional stress of combat. For example, novice soldiers could be pre-exposed to challenging VR combat stress inoculation training as has been reported by (Wiederhold & Wiederhold, 2005) with combat medics. Finally, one of the guiding principles in our development work concerns how VR can *extend* the skills of a well-trained clinician. This VR approach is not intended to be an automated treatment protocol that could be administered in a "self-help" format. The presentation of such emotionally evocative VR combat-related scenarios, while providing treatment options not possible until recently, will most likely produce therapeutic benefits when administered within the context of appropriate care via a thoughtful professional appreciation of the complexity and impact of this disorder.

Acknowledgements: The project described here has been sponsored by the Office of Naval Research, TATRC and U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

6. REFERENCES

J Difede & H Hoffman (2002), Virtual reality exposure therapy for World Trade Center Post Traumatic Stress Disorder, *Cyberpsychology and Behavior,* **5**:*6*, pp. 529-535.

DSM-IV. (1994). American Psychiatric Association, Washington, D.C.

P Gamito, C Ribeiro, L Gamito, J Pacheco, C Pablo & T Saraiva (2005), Virtual War PTSD: A methodological thread. Paper presented at the *10th Annual Cybertherapy Conference*, Basel Switzerland, June 13-17, 2005.

K Glantz, A A Rizzo & K Graap (2003), Virtual Reality for Psychotherapy: Current Reality and Future Possibilities, *Psychotherapy: Theory, Research, Practice, Training,* **40**, 1/2, PP. 55–67.

C W Hoge, C Castro, S Messer, D McGurk, D I Cotting & R L Koffman (2004), Combat Duty in Iraq and Afghanistan, Mental Health Problems, and Barriers to Care, *New England Jour. of Med.*, **351**, *1*, pp. 13-22.

L H Jaycox, E B Foa & A R Morral (1998), Influence of emotional engagement and habituation on exposure therapy for PTSD, *Journal of Consulting and Clinical Psychology,* **66,** pp. 186–192.

N Josman, E Somer, A Reisberg, A Garcia-Palacios, H Hoffman & P L Weiss (2005), Virtual reality: innovative technology for the treatment for victims of terrorist bus bombing with post-traumatic stress disorder. Paper presented at the *10th Annual Cybertherapy Conference*, Basel, Switzerland, June 13-17, 2005.

T M Keane, D G Kaloupek, E B Blanchard, F Y Hsieh, L Kolb, S P Orr, R G Thomas & P W Lavori (1998), Utility of psychophysiological measurement in the diagnosis of posttraumatic stress disorder: Results from a Department of veterans affairs cooperative study, *Jour. of Consulting and Clinical Psych.,* **66**, pp. 914-923.

N Laor, L Wolmer, Z Wiener, A Reiss, U Muller, R Weizman & S Ron (1998), The function of image control in the psychophysiology of posttraumatic stress disorder, *Journal of Traumatic Stress,* **11**, pp. 679-696.

R P Marshall, A F Jorm, D A Grayson & B I O'Toole (2000), Medical-care costs associated with posttraumatic stress disorder in Vietnam veterans, *Australian and New Zealand Journal of Psychiatry,* **34**, *6*, pp. 954-962.

A A Rizzo, M T Schultheis, K Kerns, & C Mateer (2004), Analysis of Assets for Virtual Reality Applications in Neuropsychology, *Neuropsychological Rehabilitation,* **14**, *1*, pp. 207-239.

B O Rothbaum, L Hodges, R Alarcon, D Ready, F Shahar, K Graap, J Pair, P Hebert, D Gotz, B Wills & D Baltzell (1999), Virtual reality exposure therapy for PTSD Vietnam veterans: A case study. *Journal of Traumatic Stress,* **12**, pp. 263-271.

B O Rothbaum, E A Meadows, P Resick et al. (2000), Cognitive-behavioral therapy. In: Foa, E.B., Keane, M., Friedman, M.J. (eds.), *Effective treatments for PTSD,* New York: Guilford, pp. 60–83.

B O Rothbaum, L Hodges, D Ready, K Graap & R Alarcon (2001), Virtual reality exposure therapy forVietnam veterans with posttraumatic stress disorder. *Journal of Clinical Psychiatry,* **62**, pp. 617-622.

B O Rothbaum & A C Schwartz (2002), Exposure therapy for posttraumatic stress disorder. *American Journal of Psychotherapy,* **56**, pp. 59–75.

A Treskunov, J Pair & B Swartout (2004), The Flatworld Simulation Control Architecture (FSCA): A Framework For Scalable Immersive Visualization Systems, *Proc. of the 2004 Army Science Conference*, Orlando, December 2004.

VA Fact Sheet, 12/2004, http://www1.va.gov/opa/fact/ptsd.asp Accessed on 07/03/06.

B K Wiederhold & M D Wiederhold (2004), *Virtual-Reality Therapy for Anxiety Disorders: Advances in Education and Treatment,* American Psychological Association Press: New York.

M D Wiederhold & B K Wiederhold (2005), Military mental health applications. *The 13th Annual Medicine Meets Virtual Reality Conference,* Long Beach, CA, January 29, 2005.


Figures 1, 2. *Small City Scenarios*.


Figures 3, 4. *Large City Scenarios.*

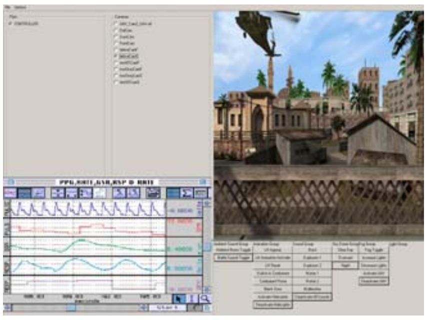

Figures 5, 6. *Desert Outskirts.*


Figures 7, 8. *Desert Roads.*


Figures 9, 10. *Building Interior ~ "Flocking" Soldier.*


Figures 11, 12. *"Flocking" Patrol ~ Turret View.*


Figures 13, 14. *Helicopter View ~ Night Vision View.*




Figures 15, 16. *Clinical Interface ~ Insurgent "Attack".*


Figures 17, 18. *Insurgent "Attack" ~ USA User Tests.*


Figures 19, 20. *User Tests – Iraq Combat Stress Control Team.*

ICDVRAT 2006

Session IV. Interaction Control

Chair: Bruno Herbelin

Designing a device to navigate in virtual environments for use by people with intellectual disabilities

P J Standen[1], D J Brown[2], N Anderton[3] and S Battersby[4]

[1,3]Division of Rehabilitation & Ageing, University of Nottingham,
B Floor, QMC, Clifton Boulevard, Nottingham NG7 2UH

[2,4]School of Computing & Mathematics, Nottingham Trent University,
Newton Building, Burton Street, Nottingham NG1 4UB

[1]*p.standen@nottingham.ac.uk*, [2]*david.brown@ntu.ac.uk*,
[3]*nicola.anderton@nottingham.ac.uk*, [4]*stephen.battersby@ntu.ac.uk*

www.nottingham.ac.uk/rehab, *www.isrg.org.uk*

ABSTRACT

As part of the process to design a device that would enable users with intellectual disabilities to navigate through virtual environments, an earlier study had collected baseline data against which to evaluate prototype design solutions. This study describes the evaluation of three design solutions: two modifications to a standard games joystick and a two handed device. Evaluation data were collected while 22 people with intellectual disabilities worked through four VE designed using a games format. None of the prototypes gave significantly improved performance over the standard joystick and some actually led to the user receiving more help from the tutor to use the device. This difference was significant for the two handed device in all four games. However there was considerable variation in results such that with some devices there was a reduction in the variability of scores between individuals. Future research needs to focus on the design of environments and how best to match the user with the device.

1. INTRODUCTION

There is a growing body of work to indicate the usefulness of virtual environments (VE) for people with intellectual disabilities (Cromby, Standen and Brown, 1996; Standen, Brown and Cromby 2001; Standen and Brown 2005; Standen and Brown in press). They have been shown to be effective in facilitating the acquisition of living skills for example shopping (Standen, Cromby and Brown, 1998) and navigating new environments in children with severe intellectual disabilities and have been developed to prepare young people for the potentially distressing experience of giving evidence in court (Laczny et al, 2001). Their three-dimensional nature allows the creation of ecologically valid settings to promote activities like choice making (Standen and Ip, 2002) which people with intellectual disabilities have limited opportunity to practice. Finally, they can provide an engaging activity for people who are frequently underoccupied and denied real world opportunities. (Standen, Lannen & Brown, 2002).

The work carried out so far has employed non-immersive VE where the environment is displayed on an ordinary computer monitor as developments could rapidly be made available for end users with limited budgets, limited technical expertise and limited backup. With the increasing power of entry level computers and the more recently commercially available platforms, the possibilities for interactive three dimensional environments have greatly increased. Developers can offer considerable choice and structure environments in order to exploit advances in educational theory (Shopland et al, 2004). There are also a variety of interfaces commercially available for interacting with three dimensional environments but these have been described as "mostly adequate…rather obstrusive and require some amount of training to use" (Bing Kang, 1998). The problems for people with intellectual disabilities in using these devices have been reported by Standen, Brown, Anderton and Battersby (in press). Using a methodology established in an earlier study they set out to systematically document the performance of a group of people with severe intellectual disabilities with the currently recommended devices (standard 3 axis games joystick or the arrow keys on a keyboard) to navigate through several interactive three dimensional environments. The results confirmed the problems reported in earlier studies (Trewin and Pain, 1999; Brown, Kerr and Crosier, 1997) and they cautioned that with problems like these, users can become frustrated and demotivated and fail to benefit from the advantages of using VE.

One solution to the problem of interfaces is to develop “natural interfaces that are intuitively simple and unobstrusive to the user” (Bing Kang, 1998). These are interfaces that have the capability to capture human gestures or biodata and translate it into code to replace standard interfaces. Bing Kang (1998) described an approach that used the orientation of the user’s face to move and orient the VE. Coyle et al (1998) have developed a non contact head controlled mouse emulator for use by quadriplegic operators to control VE; whilst Bates and Istance (2004) are developing a reliable system for eye based VE interaction. More recently, Hochberg et al (2006) implanted a 96 microelectrode array in the primary motor cortex which allowed a patient with tetraplegia to control the position of the cursor on a computer screen. These systems are more easily suited to replace a single function of the mouse either the right or left click or the control of the arrow thus limiting their usefulness in interactive educational VE. Additionally they are often difficult to calibrate and can be tiring for disabled users. Their utility will ultimately depend on how affordable they will be to a target population who also experience multiple disadvantages due to the added impact of low resources and income.

The study by Standen et al (in press) was conceived as the first step in a process to design a device that would enable users with intellectual disabilities to navigate through and interact with structured educational virtual environments. The intention was to collect information which could then be used to inform the design of a usable control device or devices and to act as a baseline against which they can be evaluated. 40 people attending a day centre for people with intellectual disabilities aged between 21 and 67 years used four environments with an equal number of sessions with the different devices being evaluated. Results indicated that first, due to the cognitive load of discriminating between the functions on one device, separate devices are retained for navigation and interaction. Secondly, resolving some of the physical difficulties with the joystick may reduce the likelihood of demotivation on initial usage and also allow better performance once use of the device has been mastered. This paper describes the evaluation of a series of design solutions for a navigation device that met the criteria set out by Brown, Kerr and Crosier (1997) that future input device design or modification should ensure that they should be operable by people with fine motor difficulties, modifiable, robust and affordable. However it was also recognised that it is highly unlikely that one device will suit all.

2. METHODS

2.1 Design

The study followed the five stage process proposed by Lannen et al (2002):

- Understand and Specify the Context of Use
- Specify User and Organisational Requirements
- Technology Review
- Produce Concept Designs and Prototypes
- Carry out a User-Based Assessment

This process is complemented by a Usability Team and a User Team. The role of the Usability Team is to guide the application of this user sensitive inclusive design methodology. The User Team should represent a cross section of the target audience and they contribute to the design process by considering design requirements and how these could be met in potential design solutions.

The earlier study (Standen et al, in press) brought the work to the beginning of the fourth stage and at this point a review of the results by the usability team decided upon two possible design solutions. The first was to modify the joystick as it allows better performance but takes time to master and secondly to devise a two handed device. This second solution had arisen from the pilot work being carried out by Lannen (2002).

The sequence of prototypes evaluated was therefore as follows:

1. Attaching a Sensitrac Flat Pad (*www.traxsys.com*) to the base of the joystick to affix it firmly to the desk. This was because during baseline evaluation, so many participants had to have the base of the joystick held to limit slipping of the device.
2. In addition to the flat pad an aluminium plate with a milled cross was fixed to the base of the joystick to limit movement to four directions (-Y, +Y, -X, +X) instead of the usual eight (additional 4 XY variants). This was because during baseline evaluation much of the help being given with the joystick

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

resulted from the participant moving in unintentional directions and because the keyboard arrows had helped in road crossing (see Figure 1).

3. Finally, a device was made which had to be controlled with both hands by mounting a Saitek wheel onto a 3 axis joystick. They were connected with a specially constructed spring loaded stock that pivoted at both ends via a ball joint to enable a floating action (see Figure 2)..

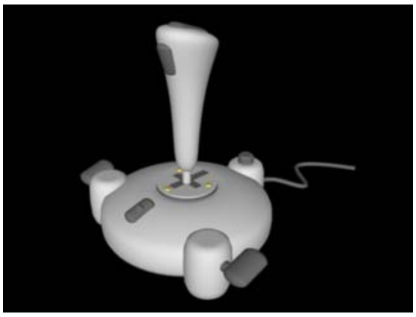


Figure 1. *A computer generated image of the joystick with the plate fixed to the base.*

Figure 2. *The two handed device.*

These were evaluated using a repeated measures design. Each member of the user team worked through four VE with each of the design solutions to enable performance data from each design solution to be compared with their own data from the currently used navigation devices.

2.2 Participants

A usability team was formed from the research team which included a computer scientist, two psychologists and a software engineer with a background in design. All worked in the area of disability. They were joined by the manager of a company that produced interfaces for people with disabilities and a manager of a day centre for people with intellectual disabilities. The user team included 22 volunteers (10 men and 12 women) aged between 21 and 67 years who regularly attended a day centre for people with intellectual disabilities and all met the requirement of having sufficient visual ability to see the VE on the computer monitor. They were selected to represent a wide range of ability within the severely disabled category. Their verbal IQ ranged from 9 to 113 and non verbal IQ from 0 to 28 so their combined scores all fell within the severely disabled range. For motor control and co-ordination, three were in the normal range, 13 showed moderate discrepancy from normal and five showed severe discrepancy with one being unable to complete the assessment.

The level of verbal ability of the user team prevented most of them from participating in group discussions so their involvement was on an individual level.

2.3 Virtual environments

Four training VE (*asteroids, dolphins, temple and road crossing* Figure 3) had been constructed in order to evaluate the currently used devices and these were once again used for this part of the study. A full description is given in Standen et al (in press). They were all designed using game format in that they consisted of varying levels of difficulty with access to each level only allowed once the correct level of performance had been achieved at the previous level and each environment constrained different possibilities in order to test a range of uses of the control devices but without presenting the user with too many options initially. The software also collected information on task achievement (scores), time taken and collisions. For two environments the device was not required to provide forward movement. Table 1 summarises the characteristics of each environment.

2.4 Data collection

Data collection took place in the day centre attended by the participants. Participants had sessions scheduled for once a week, which lasted a maximum of 30 minutes but could be terminated earlier if they wished. One of the researchers (NA) sat alongside them to give assistance and encouragement. Each session was recorded on videotape, with the camera positioned to view both the participant and the researcher sitting next to them. The videotapes were analysed using a method established in an earlier study (Standen et al, 2002) which yielded measures of help given by the researcher. For the present evaluation only the percentage of time physical assistance with navigation was used. The researcher also kept a diary to record any other information that might be useful but that would not be picked up by video analysis or the software data gatherer. Computer collected data (scores and collisions) were adjusted for length of session. Video collected data were expressed as a percentage of session duration. Statistical comparisons between the devices were made using the Wilcoxon test for paired data.

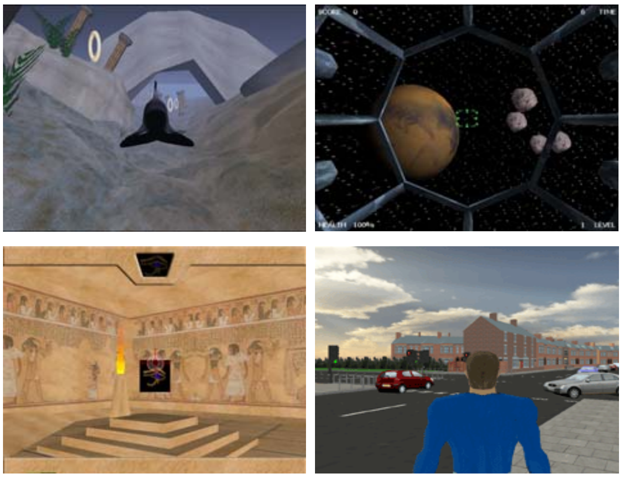


Figure 3. *Screen shots from each of the four virtual environments.*

Table 1. *Characteristics of training environments.*

	Devices used for forward movement	Devices used for up/down and left/right movement	Devices tested for interaction	Follow avatar
Asteroids	None required	joystick	joystick button	no
		mouse	LH mouse button	
Dolphins	None required	joystick	None required	yes
		arrow keys		
Temple and Road crossing	joystick	joystick	LH mouse button	yes
	arrow keys	arrow keys	LH mouse button	

3. RESULTS

Scores or collisions and percentage of session during which the tutor gave physical assistance are shown in Table 2.

Table 2. *Median adjusted scores/collisions and median percentage of session spent by tutor help.*

measures	scores		collisions	Help with navigation			
devices	dolphins	asteroids	temple	dolphins	asteroids	temple	road crossing
joystick	2.87	4.54‡	9.65	0	15.12‡	14.55‡	6.2
keyboard	2.32		8.88	0		4.22	8.41
mouse		10.64			0		
Sticky base	3.44‡*	3.57‡	8.62	11.77‡*	13.73	9.6	11.31
plate	3.18‡	3.12‡	11.44‡	8.34‡*	3.98	8.66	6.36
2 hand	3.69‡	1.93‡*	8.82	27.50‡*	33.61‡	48.67‡	27.23‡*

* Significantly different from joystick
‡ significantly different from keyboard/mouse

3.1 Addition of Sensitrac Flat Pad to the base of the joystick.

This modification did allow users to achieve higher scores in one of the environments than with the standard joystick ($p<0.017$) or the keyboard ($p<0.009$) but they were still significantly lower than those achieved with the mouse ($p<0.0003$). Where collisions were recorded there was no significant improvement and in one game users actually needed significantly more help than they did with either the original joystick ($p<0.02$) or the arrows on the keyboard ($p<0.002$). However, in two of the environments although not significant the medians for amount of help received are lower with the addition of the sticky base.

3.2 Attaching the aluminium plate

Attaching a plate to the base of the joystick that limited movement to four directions gave no improvement in scores or collisions over the addition of the sticky base. However it allowed the user to gain better scores ($p<0.05$) than they did with the arrows in the keyboard but at the same time more collisions ($p<0.05$). Medians suggest that the tutor spent less time helping users with this modification than they did with the original joystick, keyboard and sticky base but these differences did not reach significance. This lack of a significant difference is largely due to the huge variation in scores. An examination of individual scores: revealed that with the addition of the aluminium plate only 2 people received help more than 40% of time as compared to 6 people using the standard joystick.

3.3 Two handed device

On the easiest environment this prototype enabled participants to achieve higher scores than with all the other devices significantly ($p<0.005$) so when compared with the keyboard but by the time this prototype was being tested, participants were very familiar with this environment. When collisions were recorded it was no different to any of the other devices. However, with the more complicated environments participants gained fewer scores than with any of the other devices and needed more help than they did with any of the other devices. For example, significantly ($p<0.001$) more help was received than with the standard joystick in road crossing.

4. DISCUSSION

These results are disappointing as after all the consultation and testing the only clear cut result was that the two handed device, in spite of being the last one that was evaluated, was no better on performance and was worse in terms of help required to use it when compared with the standard joystick it was intended to replace. The single most effective solution was the addition of the Sensitrac flat pad to prevent the joystick base slipping. While there is no evidence in the results presented here, observations from the researcher's diary also support the use of this modification. This is a low cost solution that did not have a major impact on group results but gave many individuals confidence to experiment with the force they exerted on the joystick.

These disappointing results may be because the team was unrealistic in aiming to achieve a low cost solution: simple modifications to mass produced devices are perhaps not the way to meet the needs of this group of users. The intention to produce affordable solutions was a priority as this makes such devices available to a poorly resourced client group. Basing solutions on already available components increases the chances that the design solution would be manufactured. These requirements necessarily constrained the possible solutions to the design challenge.

In addition to the constraints imposed on the solution before the study began it was probably inappropriate to look for one solution for all the participants and also to use group results to evaluate each prototype. Although the measures recorded from each user were compared with their own measures on the other devices, an examination of individual results indicated that some individuals did much better with some prototypes than with others. This may be due to their levels of cognitive or motor ability but also to an interaction between their ability and the different requirements of each environment. For example, some participants navigated well with the cross plate attached to the joystick in the *road crossing* environment. This environment was characterised by perpendicular lines (edges of pavement or crossing) and most movement was required in straight lines with 90 degree turns. However, not all participants grasped this but those who did, performed well in this situation.

From the previous study (Standen et al, in press) it was clear that constraining the tasks for the participant (eg providing forward movement) was beneficial but this obviously constrains the situations that can be depicted in the environment and restricts the freedom to explore. In spite of these disadvantages, the immediate solution may be to focus on the design of the environments to avoid disorientation (Ruddle and Jones, 2001) and make it possible for the more disabled user to employ the easier devices such as mouse emulators. Flexible software is easier to achieve than flexible hardware.

An obvious disadvantage of an evaluation of this nature is that participants spent regular sessions throughout two years using the environments. Even allowing an initial period in which participants can become familiar with the set up and the software, their performance did not plateau and an increase in ability to play the games was taking place throughout the study. This should have led to better performance with the later prototypes whereas in fact the opposite was the case. An unforeseen advantage of this was that many participants who returned to using the standard joystick at the end of the study, appeared to have much better motor control and hand eye co-ordination. This should not be surprising given the work already reported on the use of VE for motor rehabilitation (Holden, 2005) and future work should focus on this advantage for this user group.

Acknowledgements: This research was carried out with the support of EPSRC award number GR/R21851

5. REFERENCES

R Bates, H O Istance (2004), Towards eye based virtaul environment interaction for users with high level motor disabilities. Proceedings of the 5th International Conference on Virtual Reality and Associated Technologies. Sharkey, McCrindle & Brown (eds), Oxford, UK, 18-20 Sept. 2004. pp. 275-282.

S Bing Kang (1998), Hands-free navigation in VR environments by tracking the head, International Journal of Human-Computer Studies, 48, pp. 247-266.

D J Brown, S J Kerr J Crosier (1997), *Appropriate input devices for students with learning and motor skills difficulties*: Report to the National Council for Educational Technology, UK.

E D Coyle, M Farrell, R White, B Stewart (1998), Design of a non-contact head-control mouse emulator for use by a quadriplegic operator, Proceedings of the 2nd International Conference on Virtual Reality and Associated Technologies. Sharkey, Rose & Lindstrom (eds), Skövde, Sweden, 10-11 Sept. 1998. pp. 35-43.

J J Cromby, P J Standen, D J Brown. (1996), The potentials of virtual environments in the education and training of people with learning disabilities, *Journal of Intellectual Disability Research*, **40,** 6, pp.489-501.

A Laczny, D J Brown, J Francik, P Cooke (2001), The virtual courtroom: a practical approach to virtual reality on the internet, *Studia Informatica*, **22**, 2, pp. 287-299.

T J Lannen (2002), A multi-disciplinary approach to the control of virtual environments for young people with moderate to severe learning difficulties. PhD thesis, Nottingham Trent University.

T J Lannen, D J Brown, H Powell (2002), Control of virtual environments for young people with learning difficulties, *Disability and Rehabilitation*, **24**,11-12, pp. 578-586.

L R Hochberg, M D Serruya, G M. Friehs, J A. Mukand, M Saleh, A H Caplan, A Branner, D Chen, R D Penn, J P Donoghue (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia, *Nature*, **442**, 7099, pp. 164-171.

M K Holden (2005), Virtual environments for motor rehabilitation: review, *Cyberpsychology and Behavior*, **8,** 3, pp. 187- 211.

R A Ruddle, D M Jones (2001), Movement in cluttered environments, *Presence*, **10**, 5, pp. 511-524.

N Shopland, J Lewis, D J Brown, K Dattani-Pitt (2004), Design and Evaluation of a flexible travel training environment for use in a supported employment setting, Proceedings of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies. Sharkey, McCrindle & Brown (eds), Oxford, UK, 18-20 Sept. 2004. pp. 69-76

P J Standen, D J Brown (2005), The use of virtual reality in the rehabilitation of people with intellectual disabilities, *Cyberpsychology and Behaviour*, **8**, 3, pp. 272 - 282.

P J Standen, D J Brown (in press), Virtual Reality and its role in removing the barriers that turn cognitive impairments into intellectual disability, *Journal of Virtual Reality*

P J Standen, D J Brown, N Anderton, S Battersby (in press) A systematic evaluation of current control devices used by people with intellectual disabilities in non-immersive virtual environments, *Cyberpsychology and Behavior*, **9,** 5

P J Standen, D J Brown, J J Cromby (2001), The effective employment of virtual environments in the training and rehabilitation of people with intellectual disabilities, *British Journal of Educational Technology*, **32**,3, pp. 289-299

P J Standen, D J Brown, T Proctor, M Horan (2002), How tutors assist adults with learning disabilities to use virtual environments, *Disability and Rehabilitation*, **24**, 11-12, pp. 570-577.

P J Standen, J J Cromby, D J Brown (1998), Playing for real, *Mental Health Care*, **1**, pp. 412-415.

P J Standen, W M D Ip (2002), An evaluation of the use of virtual environments in improving choice reaction time in people with severe intellectual disabilities. In: Sharkey PM, Sik Lányi C, Standen PJ, eds. *Proceedings of the Fourth International Conference on Disability, Virtual Reality and Associated Technologies.* pp. 19-24.

P J Standen, T L Lannen, D J Brown, DJ (2002), Control of virtual environments for people with intellectual disabilities. In Keates S, Langdon P, Clarkson PJ, Robinson P (eds) *Universal access and assistive technology*. Pp 63 – 72. Springer-Verlag: London

S Trewin, H Pain (1999), Keyboard and mouse errors due to motor disabilities for interaction tasks, *International Journal of Human-Computer Studies*, **50,** 2, pp. 109-144.

Tactile information transmission by apparent movement phenomenon using shape-memory alloy device

Y Mizukami and H Sawada

Department of Intelligent Mechanical Systems Engineering, Faculty of Engineering,
Kagawa University, 2217-20, Hayashi-cho, Takamatsu-city, Kagawa, 761-0369, JAPAN

sawada@eng.kagawa-u.ac.jp

www.eng.kagawa-u.ac.jp/~sawada/index_e.html

ABSTRACT

This paper introduces the development of a tactile device using a shape-memory alloy, and describes the information transmission by the higher psychological perception such as the phantom sensation and the apparent movement of the tactility. The authors paid attention to the characteristic of a shape-memory alloy formed into a thread, which changes its length according to its body temperature, and developed a vibration-generating actuator electrically driven by periodic signals generated by current control circuits, for the tactile information transmission. The size of the actuator is quite compact and the energy consumption is only 20mW. By coupling the actuators as a pair, an information transmission system was constructed for presenting the apparent movement of the tactility, to transmit quite novel sensation to a user. Based on the preliminary experiment, the parameters for the tactile information transmission were examined. Then the information transmission by the device was tested by 10 subjects, and evaluated by questionnaires. The apparent movement was especially well perceived by users as a sensation of a small object running on the skin surface or as being tapped by something, according to the well-determined signals given to the actuators. Several users reported that they perceived a novel rubbing sensation given by the AM, and we further experimented the presentation of the sensation in detail to be used as a sensory-aids tactile display for the handicapped and elderly people.

1. INTRODUCTION

Humans are able to communicate with each other by using not only verbal media but also the five senses such as vision, audition, olfaction and tactility, effectively using their body parts. Information transmitted through non-verbal media directly affects our emotions and feelings, and especially hand gestures and touch feelings play an important role for the emotional human communication. In the face-to-face communication, gestures and touch actions are effectively employed together with the information transmitted through speech and vision, and in the conversation about an actual object, the tactile sensation and haptics help to understand the object intuitively. Human communication is regarded as the information transmission through our body and sensations, and computers have recently been used as an extensive tool for supporting the communications as human-machine interfaces. Especially for supporting disabled and elderly people, computers and intelligent devices are now essential tools in the daily life.

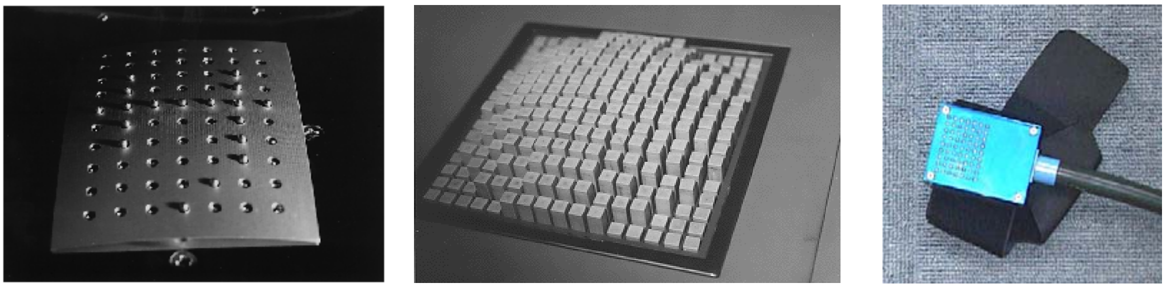
We have so far paid attention to the information transmitted through the gestural actions and the tactile sensation (Moritani et al, 2003), and are now constructing a tactile information transmission system for the presentation and communication of tactility. Tactile displays which employ piezo actuators or solenoid coils have been introduced so far (Asamura et al, 2001; Maeno, 2002; Makino et al, 2004; Moy et al, 2000; Nara et al, 2001), and some of them are commercially available now as shown in Figure 1. Such devices, however, have the problems of the requirement of high-voltage supplies and the large size of the actuators and the device itself. Furthermore, they are designed to display Braille or to present alphabetical letters and the static shape of a 2D/3D object. It is difficult to present touch feeling and tactile sensation such as a moving object and rubbing sensation, due to the restrictions caused by the problems mentioned above. This paper introduces a compact device for the presentation of tactile sensation given by the apparent movement phenomenon and the phantom sensation, by using a shape-memory alloy. A novel tactile information transmission device is developed to be employed for the tactile communication, and is evaluated by questionnaires to be used as a sensory-aids tactile display for the handicapped and elderly people.

2. VIBRATION ACTUATOR FOR TACTILE INFORMATION TRANSMISSION

2.1 Shape-memory alloy and its characteristics

A shape-memory alloy (SMA) is employed for the tactility presentation in this study. Shape-memory alloys are the metals which have the shape-memory effect. An alloy has peculiar temperature T_p, and the shape memory effect is observed when the body temperature is cooled to below T_p. In this state, it can be easily deformed, however, the original shape can be recovered by heating the body above the temperature. The effect has been widely applied in different fields such as robotic muscles, medical surgeries, actuators and elastic wires. The SMA also has a unique characteristic to shrink at a certain temperature. Figure 2 shows the schematic figure presenting the relation between the body temperature of SMA and its length. We found a specially-composed SMA whose temperatures T_1 and T_2 are low enough to be touched by human skin.

We also found a novel and interesting effect that, by making the SMA into a string, it accepts a pulse-signal to generate a vibration in accordance with the pulse frequency of the signal. Figure 3 shows a vibration actuator composed with a 5mm-long SMA string with a diameter of 0.05mm. With a weak current given to the alloy, the body temperature rises to the T_2 due to the generated heat inside the body, and it shrinks about 7% of the original length. When the current stops, the body is cooled to below T_1, to recover the original length. In this manner, by giving a pulse signal current with the frequency of several tens or hundreds hertz as presented in Figure 3, a vibration is generated to be perceived by human body as tactile sensation. The actuator has the advantage of its compactness having the diameter 3mm and the thickness 1mm, and the low energy consumption of 20mW with its quick response to generate a vibration.



(a) Vibrator array *(b) Tactile pin array* *(c) Piezoelectric actuators*

Figure 1. *Examples of tactile displays.*

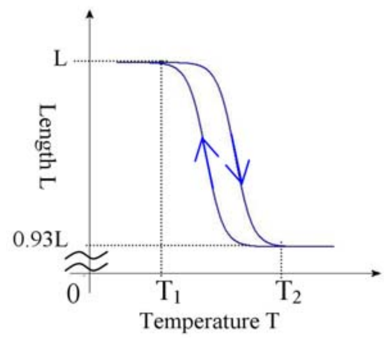


Figure 2. *Temperature characteristics of SMA.*

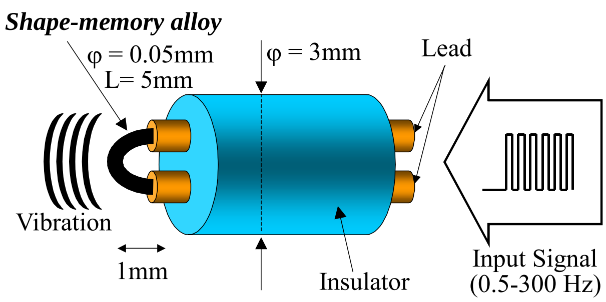


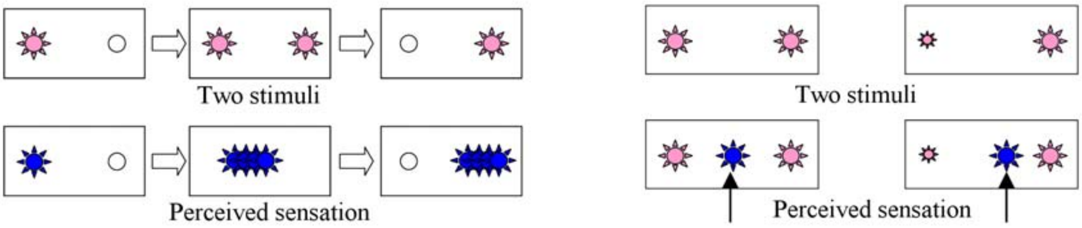
Figure 3. *Vibration Actuator.*

2.2 Higher psychological perception of tactile sensation

In this study, we paid attention to the higher psychological perception of tactile sensation for the transmission of tactile information. The apparent movement (AM) is known as one of the higher psychological perception of human tactile sensation (Bekesy, 1957). When two locations on our skin surface are excited by two mechanical vibratory stimuli with transient time delay, we perceive an illusory sensation which continuously moves from first location to the other, as shown in Figure 4 (*a*). The phantom sensation, on the other hand, is also the higher psychological perception of tactile sensation (Alles, 1970). A variable sensation appears between two locations when they are stimulated simultaneously with arbitrary intensity. If two stimuli have

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

the same intensity, the phantom sensation is perceived in the middle of them. If one stimulus is stronger than the other, the illusory sensation appears at the closer location to the stronger one, according to the strength ratio. Figure 4 (b) shows the schematic figure of the phantom sensation which appears between two mechanical stimuli.



(a) Apparent movement (AM) *(b) Phantom sensation (PS)*

Figure 4. *Higher psychological perception of tactile sensation.*

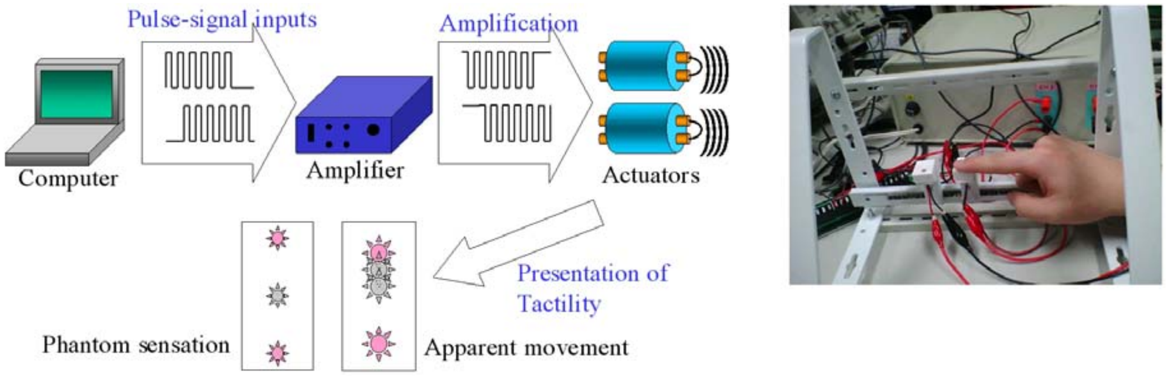


Figure 5. *Tactile information transmission system*.

2.3 Tactile information transmission using SMA actuator

The authors have developed a vibration-generating actuator electrically driven by periodic signals generated by current control circuits, for the tactile information transmission. The vibration actuator employing a SMA formed into a thread is able to provide a mechanical vibratory stimulus on a spot around 1 mm. By coupling two actuators, the PS and the AM are generated by the vibratory sensation on the skin, which are perceived as particular tactile information. Figure 5 shows the configuration of the tactile information transmission system, together with a picture of an experiment to present tactile information to a finger. A pulse-width modulated (PWM) rectangular wave signal with arbitrary frequency, amplitude and duty-ratio is generated in a PC, which is amplified to drive the two actuators. The amplifier was specially designed for driving SMA actuator in variable frequencies and variable voltage amplitude with current control. In this study, the tactile information transmission by the PS and AM were examined, and the tactile sensation by the novel device was evaluated for the development of a tactile display for the handicapped and elderly people.

3. PRELIMINALY EXPERIMENTS FOR TACTILE INFORMATION TRANSMISSION

We first conducted a preliminary experiment to find effective locations on our body for presenting tactile sensation, and also to study the optimal parameters for generating tactile stimuli by the vibration actuator. Two experiments were conducted here, which would reveal

1. the relation between body locations and the parameters of vibratory signal giving to the actuator, and
2. the optimal parameters for generating the AM.

We empirically know that the sensitivity against a mechanical stimulus differs from the location on a skin, for example, a stimulus is perceived as different intensity depending on the location of the skin surface such as palm and the back of a hand, due to the localization of the sensory cells under the skin. Two stimuli with different intensity, on the other hand, might be sensed as the same intensity when they are presented in different location of a hand. To find the relation between stimuli and body locations, the preliminary experiment was required and one subject who had a standard tactile sensitivity was employed.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

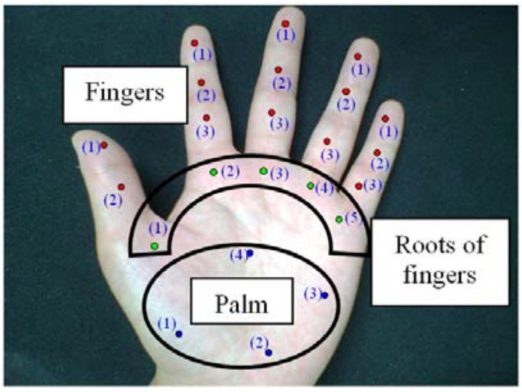


Figure 6. *Locations on palm for stimuli.*

In the experiments throughout the study, the intensity of the sensation was assessed by the scale from 1 to 10, where 10 represents the strongest sensation and 1 represents the weakest or faint to be perceived.

In the experiment 1), one vibration actuator was driven by rectangular waves with different frequencies and voltages using PWM, and the stimuli were presented to different locations on the palm of the subject's dominant hand as shown in Figure 6. By changing the frequency from 1 to 300 Hz, stimuli with the different amplitude 0.3 to 0.7 Volts were given, and the subject answered how strong he felt the tactile sensation.

Table 1. *Results of tactile sensation by the difference of location, frequency and voltage.*

Frequency [Hz]		5			50			200		
Location \ Voltage [V]		0.39	0.46	0.61	0.55	0.61	0.70	0.65	0.70	0.82
Thumb	(1)	2	3	7	3	3	4	0	1	2
	(2)	4	5	8	4	5	7	0	2	4
Index finger	(1)	1	3	5	2	2	4	0	1	3
	(2)	1	3	5	2	2	4	0	1	3
	(3)	2	4	7	2	3	5	1	3	4
Middle finger	(1)	3	4	8	2	3	4	0	0	1
	(2)	1	2	6	2	1	3	1	2	2
	(3)	2	3	7	1	2	4	1	1	1
Ring finger	(1)	2	3	6	2	3	3	1	2	3
	(2)	0	2	4	1	1	1	0	1	1
	(3)	1	3	5	0	2	2	1	2	2
Little finger	(1)	2	3	7	3	4	5	1	2	3
	(2)	1	2	6	1	1	2	1	1	2
	(3)	2	2	6	1	2	4	1	2	2
Roots of fingers	(1)	0	2	5	1	2	2	1	2	2
	(2)	2	3	4	1	5	4	1	2	2
	(3)	3	4	5	3	3	4	2	3	3
	(4)	1	1	6	2	2	1	1	1	2
	(5)	2	3	7	3	4	3	2	2	3
Palm	(1)	2	3	3	1	2	5	1	1	1
	(2)	3	4	6	1	3	6	2	2	3
	(3)	2	5	7	3	5	7	3	3	3
	(4)	0	1	2	0	2	1	1	0	1

Table 1 shows an example of the experimental result. Same numbers represent that the stimuli were perceived as the same strength of the tactile sensation. We found that the 50 Hz vibrations were highly sensitive to the tactility, and the index finger could perceive the vibratory stimuli more sensitive than the other fingers and the palm. The sensitivity of the thumb, middle and little fingers are almost the same, and higher than the ring finger. The middle part of the palm has the lowest sensitivity to the mechanical vibrations. The subject reported that he perceived the stimuli generated by different frequencies as quite different tactile sensations with each other. He felt like the beats of the blood pulse by the stimuli with the lower frequency around 1 to 5 Hz, the vibratory or tapped sensation with the frequencies around 10 to 50 Hz, and the rubbed or gently-touched sensations with the higher frequencies around 50 to 300 Hz.

We paid attention to the tapped and rubbed sensations presented by the vibratory stimuli with the frequency of 50 Hz, and tried to present the AM by using the two actuators. Two signals shown in Figure 7 are given to the actuators, which are pressed against two points of the surface of the skin, and the experiment 2) for the generation of the AM was conducted. Figure 8 shows the 15 points of the stimuli presentations, and two stimuli are presented by two actuators to arise the AM between arbitrarily selected two locations. In this experiment, the transient time delay B [msec] and the amplitude of the signal H [V] shown in Figure 7 are changed, and the subject answered the perception of the AM.

An example of the experimental result is shown in Table 2. The AM appeared with the time delay of 100 to 800 msec and the amplitude of 0.6 to 0.85 volts, on all the locations. Among the conditions, the index and middle fingers have the higher sensitivity to the AM sensation with the time delay of 500 msec. The root of the finger and the center of the palm also have comparatively high sensitivity. When the time delay is smaller than 50 msec, the two stimuli reinforce each other and one phantom tactile image appears in the middle of the two stimuli to work to perceive the phantom sensation.

Based on the two experiments conducted by the subject, the experimental conditions for the AM presentations were established. The frequency 50 Hz and the amplitude 0.75 volts of the input signal for the actuators were set as the basement, and the index and middle fingers and the palm were selected as the AM presentation locations.

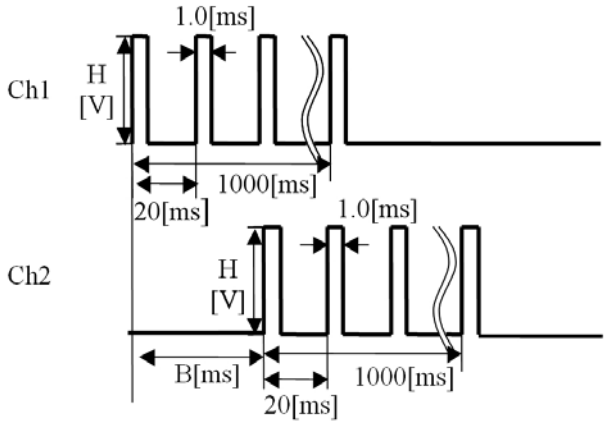


Figure 7. *Signal for AM generation (50Hz).*

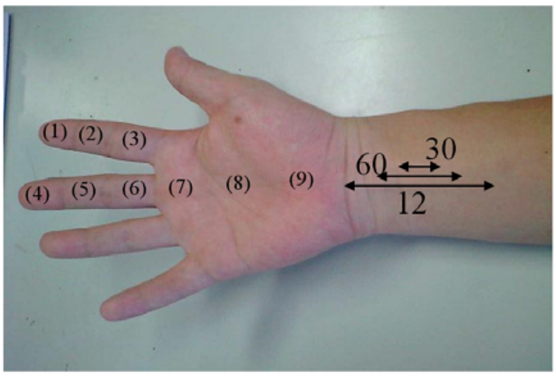


Figure 8. *Presentation points in AM evaluation.*

Table 2. *AM perception by the difference of location, frequency and time delay.*

Time delay [ms]		50			500			1000		
Location \ Voltage [V]		0.70	0.75	0.82	0.70	0.75	0.82	0.70	0.75	0.82
Index finger	(1)−(2)	0	0	1	2	2	3	1	3	2
	(2)−(3)	0	1	1	2	3	3	2	3	2
	(1)−(3)	1	1	0	2	3	4	1	1	1
Palm	(7)−(8)	2	2	2	2	2	2	0	1	0
	(8)−(9)	1	2	1	1	1	2	0	1	1
	(7)−(9)	1	1	1	1	2	1	1	2	1
Wrist	30	1	1	0	1	1	1	1	1	1
	60	1	1	0	1	1	2	2	2	3
	120	0	0	0	0	2	1	1	1	1

4. EXPERIMENTS FOR AM PRESENTATION

4.1 Presentation of the apparent movement

The presentation of the AM sensation given by the actuators was evaluated by 10 subjects, based on the experimental conditions determined by the preliminary experiments. The time delay was fixed to 500 msec, and the explanation of the apparent movement of tactile sensation was given to the subjects in advance. In the experiment, two AM stimuli, which run in two directions, one from the tip to the root of a finger and the other from the root to the tip, were randomly presented to the subjects, and they answered the direction. The trial repeated five times by generating the AM stimuli with various signal parameters, and questionnaires were conducted to the subjects after the experiment.

The result of the experiment is shown in Table 3. The numbers in the table represent the correct answers out of five trials. The AM sensation given by the actuators was perceived perfectly by all the subjects, and the moving direction was recognized 97 %. Several users reported that a rubbed sensation or a small object running across the palm was distinctively perceived by the AM presentation. All the subjects commented that the sensation was quite novel and they could clearly recognize the trajectory of the phantom motion in their hands.

4.2 Evaluation of rubbed sensation

The experiments described above showed that the AM stimuli generated by the vibration actuators were effectively used for the tactile information transmission. Several subjects reported that they felt something rubbing on the finger and the palm, and also they sensed a small virtual object moving on the palm. We also found that this novel sensation was perceived only in particular conditions of the vibration patterns of the actuators. It depends specially on the transient time delay between two vibratory stimuli generated by the actuators. Another experiment was conducted to find the optimal time delay to present the rubbed sensation on a skin.

Table 3. *Result of AM presentation experiment (Number of correct answers out of five trials).*

Location / Subject	Index finger		Middle finger		Palm	
	(1)–(2)	(1)–(3)	(4)–(5)	(4)–(6)	(7)–(8)	Half
A	5	5	5	5	5	5
B	5	5	4	5	5	4
C	4	5	4	5	5	5
D	5	3	4	5	5	5
E	5	5	5	5	5	5
F	5	5	5	5	5	5
G	5	5	5	5	5	4
H	5	5	5	5	5	5
I	5	5	5	5	5	5
J	5	5	5	5	4	5
K	5	5	5	5	5	5

Table 4. *Result of the presentation of rubbing sensation.*

Time delay [ms] / Subject	100	200	300	400	500	600	700	800	900	1000
A	–	–	–	O	O	O	O	–	–	–
B	O	O	O	O	O	O	O	O	–	–
C	O	O	O	O	O	O	O	O	–	–
D	O	O	O	O	O	O	O	O	–	–
E	O	O	O	O	O	O	O	O	–	–

The vibratory stimuli, whose time delay was changed from 100 to 1000 msec, were given to 10 subjects, and they answered whether they could perceive the rubbed sensation. Two locations for the presentation were (1) and (3) of the index finger shown in Figure 8, and the frequency and the amplitude of PWM signal were set to 50 Hz and 0.75 volts, respectively.

Table 4 shows the results of the experiment obtained by the subjects A to E. Almost all the subjects perceived the sensation with the time delay 100 to 700 msec, and they asserted that the moving speed of the sensation changes according to the time delay. With no time delay, phantom sensation appears between two vibratory stimuli, and at 100 msec the sensation instantaneously moves from the one to the other. As the time delay increases, the moving speed decreases, and at around 800 msec, the sensation cannot be perceived. Some subjects claimed that with the delay around 700 and 800, the sensation started moving from the one stimulus, and disappeared in the middle between the two stimuli, and then appeared again to move to the second stimulus. This phenomenon is quite interesting, and should be further examined in the future studies.

With the experiment, we confirmed that the novel sensation was effectively given by the vibration actuators, and the rubbed sensation would be generated with the time delay around 500 msec.

5. TACTILE INFORMATION TRANSMISSION BY VIBRATION ACTUATORS

For the tactile information transmission, we are paying attention to the AM and the rubbed sensation to realize a sensory-aids tactile display for the handicapped and elderly people. By using the phantom sensation, a tactile phantom image appears in any location between two vibratory stimuli. Furthermore, with the apparent movement, one can perceive a sensation moving from one location to the other, under the control of the time delay between two vibratory stimuli. The moving speed can be manipulated by controlling the parameters to drive the actuators.

We designed and constructed a device for the tactile information transmission by arranging 9 actuators two-dimensionally on a flat plate as shown in Figure 9. The device consists of 3 x 3 array of actuators, which are independently driven by nine amplifiers. We held an experiment to present simple letters on a palm by showing the order of writing. This would be realized by employing the AM as shown in Figure 10. By driving the actuators in the order shown by arrows, the letters were successfully perceived and preferably recognized by the subjects. All the subjects also answered that the device presented the letters much clearly than the conventional tactile display which gives a static pattern of a letter.

Figure 9. *3 × 3 array of vibration actuators.*

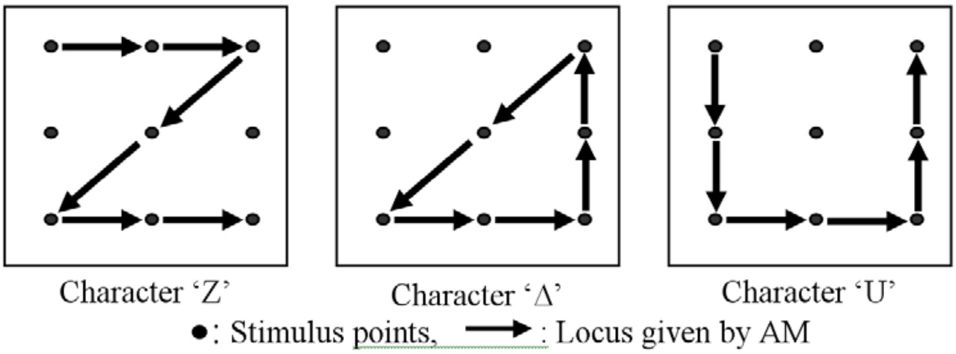


Figure 10. *Presentation of writing motion by AM.*

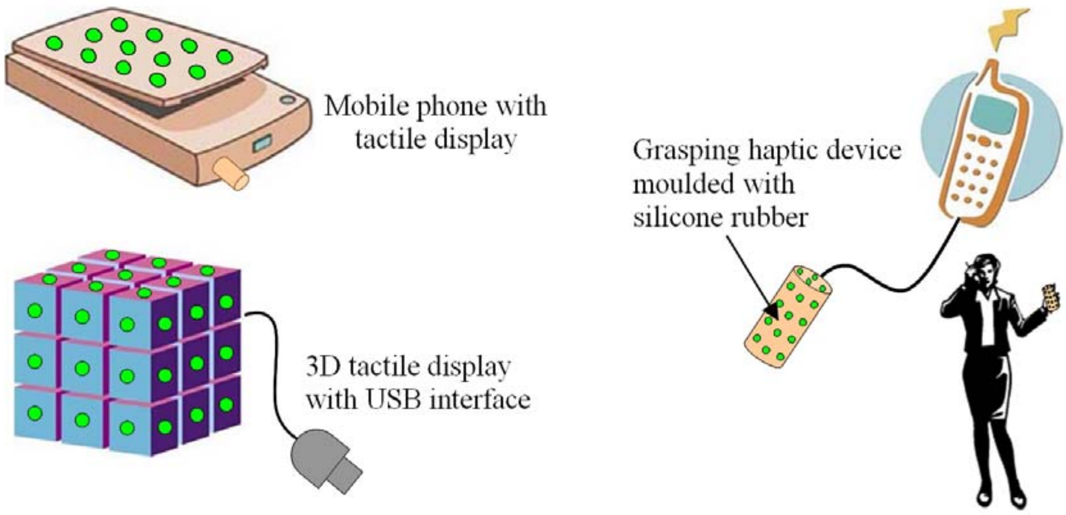


Figure 11. *Ideas for the tactile device.*

6. CONCLUSIONS

This paper introduced a compact device for the presentation of tactile sensation given by the apparent movement phenomenon and the phantom sensation, by using a shape-memory alloy. The vibration-generating actuator electrically driven by PWM rectangular signals had the advantage of its compactness, low energy consumption and quick response, and was effectively used for the presentation of novel tactile information. The device had been developed to be employed for the tactile communication, and was evaluated by questionnaires to be used as a sensory-aids tactile display for the handicapped.

The experiments showed that the device successfully presented the apparent movement and the rubbed sensation, which was preferably accepted by all subjects. Based on the evaluation results, the novel tactile display with the 3 x 3 arrangements of vibration actuators was developed, and was experimentally employed for the alphabetic character presentation on a palm by showing the order of writing. The presentation was also clearly recognized and preferably accepted. The results proved and promised the potential of the actuator using the shape-memory alloy for the tactile information transmission.

The authors are now developing novel tactile devices for the handicapped and elderly people to assist the information presentation and communication employing the tactile and haptic sensation. Figure 11 shows three examples of ideas of the tactile and haptic devices. Owing to the low energy consumption and its thin thickness, the device can be directly installed on an electric circuit board to present tactile information.

7. REFERENCES

D. S. Alles, Information Transmission by Phantom Sensations, *IEEE Trans. MMS*, 11, pp.85-91, 1970.

N. Asamura, T. Shinohara, Y. Tojo, N. Koshida and H. Shinoda, Necessary Spatial Resolution for Realistic Tactile Feeling Display, *IEEE International Conf. on Robotics and Automation*, pp. 1851-1856, 2001

G. V. Bekesy, Sensations on the Skin Similar to Directional Hearing, Beats, and Harmonics of the Ear, *J. Accoust. Soc. Amer.*, 29, pp. 489-502, 1957.

R. Maeno, Tactile Sensation of Humans -Mechanical Dynamics, Analogy and Illusion, *IEE J. Sensors and Micromachines Society*, Vol. 122-E, No. 10, pp. 469-473, 2002

Y. Makino, N. Asamura and H. Shinoda, A Whole Palm Tactile Display Using Suction Pressure, *IEEE International Conference on Robotics and Automation*, pp. 1524-1529, 2004

G. Moy, U. Singh, E. Tan and R. S. Fearing, Human Psychophysics for Teletaction System Design, *Haptics-e The Electronics Journal of Haptics Research*, Vol. 1, No. 3, 2000

H. Moritani, Y. Kawai and H. Sawada, Intuitive Manipulation of a Haptic Monitor for the Gestural Human-computer Interaction, *International Workshop on Gesture and Sign-Language based Human-Computer Interaction*, pp.5-, 2003

T. Nara, M. Takasaki, T. Maeda, T. Higuchi, S. Ando and S. Tachi, Surface Acoustic Wave Tactile Display, *IEEE Computer Graphics and Applications*, Vol.21, No.6, pp.56-63, 2001

Statistical estimation of user's intentions from motion impaired cursor use data

P Langdon[1], S Godsill[2] and P J Clarkson[3]

[1,3]Engineering Design Centre/[2]Signal Processing and Communications Laboratory,
Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, UK

[1]*pml24@eng.cam.ac.uk*, [3]*pjc10@eng.cam.ac.uk*, [2]*sjg@eng.cam.ac.uk*

[1,3]*www-edc.eng.cam.ac.uk*. [2]*www-sigproc.eng.cam.ac.uk*

ABSTRACT

We report the application of new statistical state space filtering techniques to cursor movement data collected from motion impaired computer users performing a standard Fitts's Law style selection task. Developed as an alternative to expensive haptic feedback assistance, the aim was to assess the feasibility of the basic techniques in resolving the users intended trajectory from the extremely variable and wavering data that result from the effects of muscular spasm, weakness and tremor. The results, using a choice of basic parameter for the filters, show that the state space filtering techniques are well suited to estimating the intended trajectory of the cursor even under conditions of extreme deviation from the direct track and that these filters effectively act as an extreme cursor smoothing system. We conclude that further development of the approach may lead to more effective adaptive systems capable of providing smoothed feedback to the user and estimates of intended destination. A similar approach might further be applied to situationally induced movement perturbations.

1. INTRODUCTION

Current computer input systems are often difficult for users with motion impairment to access. Impairments can result from athetoid, ataxic and spastic Cerebral Palsy, Muscular Dystrophy, Friedrich's Ataxia, Tetraplegia, spinal injuries or disorder, Parkinson's disease, stroke and arthritis. Frequent symptoms include tremor, spasm, poor co-ordination, restricted movement, and reduced muscle strength. Computers offer greater freedom to participate in education and leisure activities, as well as increased job potential and satisfaction. The Internet is a prime example that offers a great opportunity for disabled users (Nelson, 1994) enabling dialogue that is entirely independent of the ability to speak clearly, and eliminates prejudices based on appearances. We investigate ways in which computer access for people with such conditions may be improved through the use of software based cursor trajectory filtering and smoothing. Previous work in Computer Access for Motion Impaired Users we have carried out, aimed to characterise motion-impaired users and test methods for making computers more accessible, such as force-feedback. The aim of the present work is to focus on an alternative approach, that of software modification of impaired cursor motions using some of the original unassisted cursor movement data.

2. BACKGROUND

The available body of academic theory in HCI and user modelling concentrates almost exclusively on able-bodied users. There are a number of original input approaches such as head movement input (LoPresti00) and EEG scanning (Moore, 2000), but these are often expensive and frequently do not match the typical actions required to interact with a GUI (Sears, 2003). An approach that is complementary to the user's preferred input device may be more productive. One approach that offers significant potential is haptic feedback, particularly force feedback with easily available low-cost haptic devices (Dennerlein, 2000, Langdon, 2002). However, this requires the use of a mechanical, electromagnetic device to generate the instantaneous forces required. A more inexpensive solution is visual feedback resulting from software assisted cursor smoothing tested with motion-impaired movement.

As a first step to investigating this possibility we have analysed some motion-impaired cursor movement data from previous studies with haptic feedback systems. To pursue this, unassisted cursor movement data was derived from our previous work and used to test the effectiveness of advanced techniques for enhancement derived from them, namely state-space filtering. The aim was to develop and evaluate software methods of enhancing situational and health induced motion-impaired movement for dissemination to the motion-impaired and software development communities.

2.1 Force Feedback Studies

This paper builds on previous work done under the original Computer Access for Motion Impaired Users project, whose aim was to characterise motion-impaired users and test methods for making computers more accessible, such as force-feedback. Techniques aimed at enhancing cursor movement for motion impairment will also impact cursor use for able-bodied users during conditions of situationally induced perturbation or vibration. Such conditions are encountered, for example, in the use of mobile devices on vehicular transport. Current interface design practices are based on user models and descriptions derived exclusively from studies of able-bodied users. However, such users are only one point on a wide and varied scale of physical capabilities. The overall aim of this research is to contribute to the enhancement of accessible input systems and interfaces. Our previous CAMIU research (e.g. Hwang, 2002, 2005) has shown that there are very important differences between those with motion-impairments, be they elderly or disabled, and able-bodied users when they interact with computers, and that motion impairment may be ameliorated by force feedback ("Haptics": See Figure 1). In addition, ordinary users may be disabled by environmental factors. The attentional, perceptual and physical demand of mobile interactions gives rise to a Situationally Induced Impairment and Disability (SIID), contrasted with Health Induced Impairment and Disability (HIID) (Sears, 2003). We focus on investigations of software-based enhancements of cursor movement for all motion-impaired users, particularly those who are health impaired.

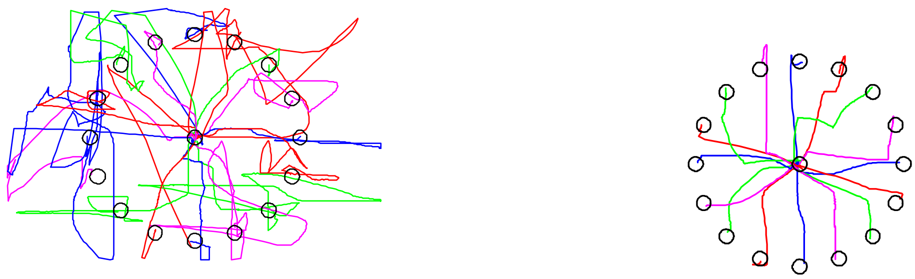


Figure 1. *Cursor Movements of a Motion Impaired user to select a target in an undamped (left) and force-feedback damped (right) condition.*

2.2 Pointing Task

Performance in pointing tasks is most commonly evaluated using speed and accuracy. Although traditional measures may show that a difference exists between conditions for impaired movement, establishing why they exist is more likely to be accomplished by analyzing the path of movement throughout a trial in conjunction with conventional measures (Langdon, 2002, Hwang, 2004). Users may trade speed against accuracy, achieving an optimised number of sub-movements that is often about two with able-bodied subjects (Meyer, 1990). Such a model may also be applicable to the analysis of haptically modified cursor movement where impaired motion can be characterised as consisting of multiple slow trajectory jumps separated by breaks during which users assess their own performance under conditions of perturbation from their motor system (Hwang, 2005).

2.3 State Space Trajectory Estimation

For several decades researchers have studied the estimation of state trajectories in state-space or dynamical models (`filtering'), using Bayesian filtering methodologies such as the classical Kalman filter. In recent years, the massive increases in available computational power have led to rapid advances in both tracking methodology and applications; in particular, the development of Sequential Monto Carlo (SMC) methods that are readily able to handle non-linear, non-Gaussian models in a vast range of applications from audio signal processing, through to computer vision and robotics (Liu, 1998, Doucet, 2000, Godsill, 2004). These methods are particularly strong candidates for the cursor movement application, since they automatically store a randomised range of hypotheses about the tracking parameters when the data are ambiguous (See Figure 2).

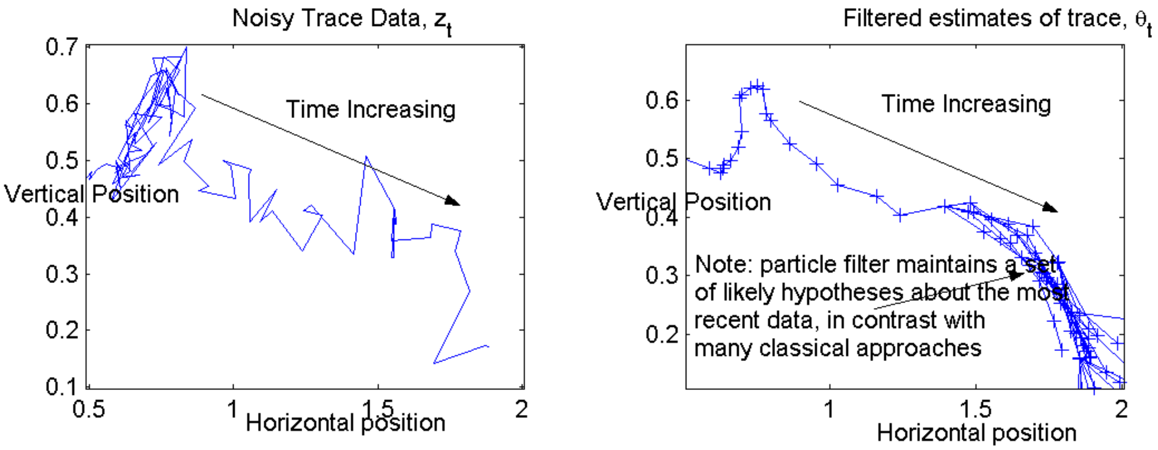


Figure 2. *Left hand side: Noisy trace data z_t generated from a near-constant velocity model in Gaussian noise (range/bearing domain). Right hand side: estimated paths $\boldsymbol{\theta}_t$ using variable rate particle filter. (Godsill04b). Note the models automatically adapt in the presence of 'tremor' like behaviour.*

A state-space modelling approach provides a methodology to explore the determination of cursor characteristics by adaptation of model, or 'state' parameters. This can be carried out in conjunction with experimental determination of the effectiveness of 'filtered' performance, in particular, improvements over no-assistance and force-feedback conditions and under conditions of situational impairment such as vibration. The results of the initial studies can then be used to arrive at an algorithm capable of assessing the extent to which modification of the filtering parameters alone is adequate before the interface elements themselves need to be modified. In the most basic state-space modelling approach, we wish to estimate a user's desired cursor position at a time **t**, say $\boldsymbol{\theta}_t$. In the simplest setting, the user's next move to $\boldsymbol{\theta}_{t+1}$ may depend only upon the current position $\boldsymbol{\theta}_t$, (which would be an x-y coordinate) through a conditional probability law, $f(\boldsymbol{\theta}_{t+1} \mid \boldsymbol{\theta}_t)$. Appropriate models for this might be the nearly-constant velocity laws applied commonly in radar tracking (Bar-Shalom, 1988). We then need to specify an observation of the move the user actually makes, given his intended move to $\boldsymbol{\theta}_{t+1}$. This again will be a random process, depending on the type of motion-impairment and/or environmental disturbances present. The probability distribution of the new measured position $\mathbf{z}_{t+1}$ might in the simplest case depend only on the desired position $\boldsymbol{\theta}_{t+1}$, through a probability density function $g(.|.)$, so that the entire system could be summarised in standard probabilistic state-space form as follows: -

$$\boldsymbol{\theta}_{t+1} \sim f(\boldsymbol{\theta}_{t+1} \mid \boldsymbol{\theta}_t) \tag{1}$$

$$\mathbf{z}_{t+1} \sim g(\mathbf{z}_{t+1} \mid \boldsymbol{\theta}_{t+1}) \tag{2}$$

Once appropriate distributions are specified for f () and g () above, Bayesian filtering can be applied to the above model in order to estimate the desired position z_t.

3. RESULTS

As a baseline test, we adapt a class of tracking models first presented in Godsill and Vermaak (2004,2005), in which the user's intended cursor path obeys a dynamic model expressed in terms of its instantaneous heading angle in the x-y plane of the monitor and the distance moved along the path. The desired trajectory is modelled as a point mass moving in a plane in a viscous medium, and subject to random forcing parallel and perpendicular to its path. The forcing is piecewise constant relative to the heading angle and subject to random change at random time instants. Such a model has been found appropriate for manoeuvring craft in tracking applications, and here we adopt it for its flexible characteristics and easy implementation within a SMC framework. This model, known as the intrinsic coordinate model, specifies the form of the dynamic model $f(\boldsymbol{\theta}_{t+1} \mid \boldsymbol{\theta}_t)$ for the user's desired cursor trajectory. We now specify a model for the user's observed behaviour, conditional upon their desired path. We adopt a very simple behavioural model that assumes the user has visual feedback of where the mouse trace is currently placed, z_t , and where the desired position should be, $\boldsymbol{\theta}_t$. The user then attempts to move the cursor to the desired position, but makes a random error, both in magnitude and direction of the move. The assumed statistics of the error are used to specify the likelihood function $g(z_{t+1}|\ \boldsymbol{\theta}_{t+1})$. Having specified the dynamical model for the desired cursor trajectory and the observation (dynamical) model for the actual path followed, we may then infer the desired path using a Bayesian filter, i.e. we obtain recursively an estimate of the 'filtering distribution', $p(\boldsymbol{\theta}_t|z_{1:t})$. This is implemented using a special version of the sequential Monte Carlo filter, the variable rate filter (Godsill and Vermaak (2004,2005)).

Some example results from this setup are given in Figure 3 below, using three of the motion impaired trajectories from the set illustrated in Figure 1.

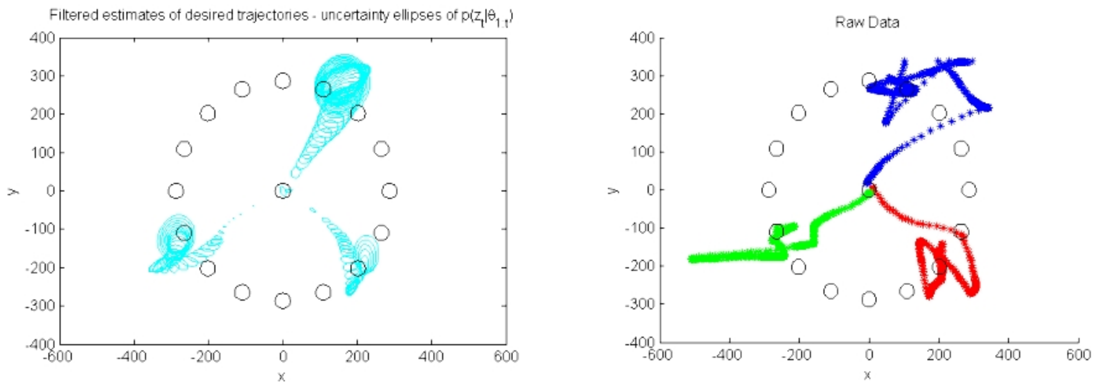


Figure 3. *Left hand side: filtered traces, showing confidence ellipses obtained from the sequential Monte Carlo filter. Right hand side: corresponding noisy trace data. Note that the filter has 'smoothed' much of the erratic jumping behaviour and quite clearly identified the user's intended destination, despite the ambiguity of the data.*

4. DISCUSSION AND CONCLUSIONS

This approach is an innovation in enabling access to computers for motion-impaired users that has not been researched in depth before. Previous work with able-bodied interface use has looked at trajectory analysis for prediction of the users' intended target (e.g. Oirshot, 2001) and trajectory prediction for telepointers under conditions of network latency delays (e.g. Gutwin, 2003). In the former approach, able-bodied trajectories proved to be essentially linear and velocity predictable while in the latter, encouragingly, trajectories for writing-like gestures were successful predicted up to delays of 80 msec and could be smoothed using Kalman filter predictors. Both able-bodied physiological tremor and pathological tremor, common in Parkinson's disease, have been addressed using Fourier software methods (Riviere96) with some success, indicating that software approaches are feasible for predictable motion-impairments with a repetitive oscillatory nature.

We have demonstrated the utility of the state space filtering approach with results from preliminary analysis of a range of motion impaired cursor movement traces taken under different conditions and proved the effectiveness of the technique for further development.

Following these findings, new models non-linear/non-Gaussian models can be proposed and tested using state-of-the-art sequential Monte Carlo methods. Further key developments that are possible include:-

1. **The effect of visual (and other) feedback**: there are many options for feedback of the filtering results to the user, depending on the extent to which the given (software) application is modified. For example, the estimated intended cursor position $\boldsymbol{\theta}_t$ could be displayed on screen in place of the actual physical cursor position $\mathbf{z}_t$, thus introducing elements of feedback into to the system;
2. **Direct determination of user's intentionality:** If it is possible to access to the configuration of the screen in an application then it will be possible to formulate and solve a more directed approach to user intentionality. One would then have a collection of dynamical models $f(\boldsymbol{\theta}_{t+1}\ |\ \boldsymbol{\theta}_t, \mathrm{D})$, one for each possible destination D, which can be built explicitly into the tracking framework;

3. **Automated parameter adaptation (unsupervised learning):** It might be possible to first model a number of different classes of behaviour and then estimate the correct class for a given user type. The class variable could be estimated as another unknown state of the system. Alternatively, a generic model could be specified that would then adapt automatically as part of the state-space model. In either case, improved performance could be expected, as the system would automatically adapt to the different characteristics of users.

4.1. Implications

The preliminary results suggest that State Space Trajectory Estimation will permit a new way of using cursors under conditions of induced impairment. Studying the properties of these algorithms interacting with specific induced impairments will extend the available academic knowledge about HIID or SIID motion-impaired user interaction with computers and furthermore develop new methods for making computer use more accessible. This potentially could be a bolt-on addition to existing GUIs, and guidance for user-interface development kits, so will be directly usable.

If successful, software based cursor analysis will benefit those whose motion-impairments inhibit the use of a keyboard and mouse for interacting with a computer. Those who already have limited access will benefit from easier computer access and those who are at present too severely impaired to use computers may be enabled to do so for the first time. This will provide them with an empowerment not previously experienced, from the ability to be more independent, to new vocational opportunities. It seems likely that users of diverse mobile computing equipment could benefit from the facility to selectively filter out unwanted effects of moving environments, for example, working on a train, without compromising system response. The results will also be applicable to users whose interaction is limited through natural ageing or deficiencies in the physical interface.

Acknowledgements: This work was sponsored by the UK Engineering and Physical Sciences Research Council and the University of Cambridge.

5. REFERENCES

Dennerlein J.T. *et al.* (2000) Force feedback improves performance for steering and combined steering-targeting tasks, *Proceedings of ACM SIGCHI CHI 2000*, 423-429.

Doucet, A. and Godsill, S.J. and Andrieu, C. (2000), On sequential Monte Carlo sampling methods for Bayesian filtering, Statistics and Computing,2000, v10, pp-197-208.

Godsill, S.J., and Doucet, A. and West, M. (2004), Monte Carlo Smoothing for nonlinear time series. Journal of the American Statistical Association. Vol.50, pp. 438-449,.

Godsill, S.J. and Vermaak, J. (2004) Models and algorithms for tracking using trans-dimensional sequential Monte Carlo. In Proc. IEEE ICASSP.

Godsill, S.J. and Vermaak, J. (2005) Variable rate particle filters for tracking applications. In *Proc. IEEE Stat. Sig. Proc., Bordeaux.*

Gutwin, C., Dyck, J., Burkitt, J., (2003), Using Cursor Prediction to Smooth Telepointer Jitter, ACM.SIGGROUP, Supporting Group Work, pp294-301,.

Hwang, F., Keates, S., Langdon, P.M., Clarkson, P.J., Mouse Movements of Motion-Impaired Users: A sub-movement analysis, ACM ASSETS 2004.

Hwang, F., Langdon, P.M., Keates, S., Clarkson, P.J. and Robinson, P. (2002) 'Cursor characterisation and haptic interfaces for motion-impaired users' in Universal Access and Assistive Technology, 1, 87-96.

Hwang, F., Keates, S., Langdon, P. And Clarkson, P.J. (2005) 'A submovement analysis of cursor trajectories' in *Behaviour and Information Technology*, **24** (3), 205 - 217 (Spring)

Langdon, P.M., Hwang, F., Keates, S., Clarkson, P.J. Robinson, P. (2002) 'Developing assistive interfaces for motion-impaired users using cursor movement analysis in conjunction with haptic feedback' in *(ICDVRAT) 2002, Veszprem, Hungary*, 223-230.

Liu, J S and Chen, R. (1998), Sequential Monte Carlo methods for dynamical systems, JASA, , 93, pp1032—44.

LoPresti E. (2000) Neck range of motion and use of computer head controls, *ACM ASSETS 2000*, 121-128.

Meyer, D.E., Abrams, R.A., and Wright, C.E., (1990), Speed-accuracy tradeoffs in aimed movements: toward a theory of rapid voluntary action. In M. Jeannerod, (Ed.), *Attention and Performance XIIII*, Hillsdale, NJ: Lawrence Erlbaum 173-226.

Moore M.M. and Kennedy P.R., (2000), Human factors issues in the neural signals direct brain-computer interface, *Proceedings of ACM ASSETS 2000*, 114-120.

Nelson J.A. (1994) The virtual community: A place for the no longer disabled, *Proceedings of the 2nd Annual International Conference on Virtual Reality and Persons With Disabilities*, Northridge, USA, pp. 98-102.

Oirschot, H.K., Houtsma, J.M. Cursor Trajectory Analysis, in Brewseter, S., Murray-Smith, R., (Eds.) Computer Interaction, Springer, LNCS 2058, pp127-134.

Riviere, C.N., Thakor, N.V., (1996), Modeling and Canceling Tremor in Human Machine Interfaces, IEEE Engineering in Medicine and Biology, pp29-36.

Sears, A., Lin, M. , Karimullah, A.S. (2002), Speech-based cursor control: understanding the effects of target size, cursor speed, and command selection, Universal Access in the Information Society Vol 2- 1, pp30-43., Springer-Verlag.

Sears, A. Karat, Oseitutu, K. Karimullah, A. and Feng, J. (2001) Productivity, satisfaction, and interaction strategies of individuals with spinal cord injuries and traditional users interacting with speech recognition software, Universal Access in the Information Society, 1, 1-12.

Sears, A., Lin, M., Jacko, J. and Xiao, Y. (2003) When computers fade… Pervasive computing and situationally induced impairments and disabilities. Proceedings HCI International 2003, pp. 1298-1302.

Y. Bar-Shalom, Y. and Fortmann, T.E.,(1988), *Tracking and Data Association*. Academic Press,.

Hands-free man-machine interface device using tooth-touch sound for disabled persons

K Kuzume and T Morimoto

Department of Information Engineering, Yuge National College of Technology
1000 Yuge, Kamishima-cho, Ochi-gun, Ehime Ken, JAPAN

kuzume@info.yuge.ac.jp, ap05007@yuge.ac.jp

ABSTRACT

This paper presents the realization of a hands-free man-machine interface using tooth-touch sound. The proposed device has several advantages, including low price, ease of handling, and reliability. It may be used as an Environmental Control System (ECS) and communication aid for disabled persons. We analyzed the characteristics of the tooth-touch sound, obtained using a bone conduction microphone. We then designed the device using VHDL (Hardware Description Language) and a simulation of the FPGA (Field Programmable Device) in practice. We applied our device to the ECS to demonstrate its usefulness and evaluate its performance. The results confirmed that the proposed device had superior features to comparable devices, such as those utilizing voice control or eye blinks, chin operated control sticks, mouth sticks, or a brain computer interface (BCI) for severely disable persons.

1. INTRODUCTION

Severely disabled persons have proven to be at risk of negative life indicators, such as a high risk of depression and anxiety (P. MacIsaac, Ashley Craing, et. al. 2002). Effective aids are necessary to improve their quality of life and lower risks. Viable therapies include psychological, medical or pharmacological, physical and occupational therapies. None-the-less, reducing dependency is a desirable move towards improving the quality of life of the disabled. This may be achieved through the development of technological aids to enhance their ability to control devices in the environment.

Environmental Control Systems (ECS) enable disabled persons to activate and control their environment. ECS devices should improve the quality of life for disabled persons and lessen the burden on helpers. In the last ten years, thanks to progress in information technology, many types of ECS devices have been developed. Some resemble the remote control units used with television sets, making use of finger movements for operation. Unfortunately, many disabled persons have limited or no such control from the neck down. A number of devices have been developed to meet the needs of the severely disabled. These include devices utilizing suck-puff techniques, voice control, eye blinks, chin operated control sticks and mouth sticks (M. P. Barnes 1994).

In recent years, a new ECS, called the Mind Switch (MS), has been developed. It makes use of variations in the electroencephalograph (EEG) patterns formed when people intend to control the environment. The MS is mounted in a plastic housing and attached to a moisture repellent cap. To make contact with the head, commercially available electrode gel or paste is used. The disabled must spend considerable time learning how to control their EEG in order to communicate or activate devices (A. Craing, P. MaIsaac, et. al. 1997). Moreover, as the EEG signal is very small, the amplitude is easily disrupted by background noises. As such, sophisticated noise suppression techniques are necessary. (L. Kirkup, A. Searle, A. Craig, et. al. 1997)

In this paper, we propose a "hands-free man-machine interface device utilizing the tooth-touch sound". The tooth-touch sound signal is detected by a bone conduction microphone, after which it can be easily processed. The proposed device is superior to conventional devices in the following ways:

- Low price
- Fitness
- Ease of handling
- Reliability

This paper is organized as follows. In Section 2, we analyze the characteristics of tooth-touch sounds. Section 3 is devoted to design of the interface's device architecture and evaluation of its performance through computer simulation. Section 4 applies our device in the environmental control system of disabled persons. Section 5 outlines our conclusions and potential development.

2. ANALYSIS OF TOOTH-TOUCH SOUND SIGNAL

2.1 Measurement of the tooth-touch sound signal

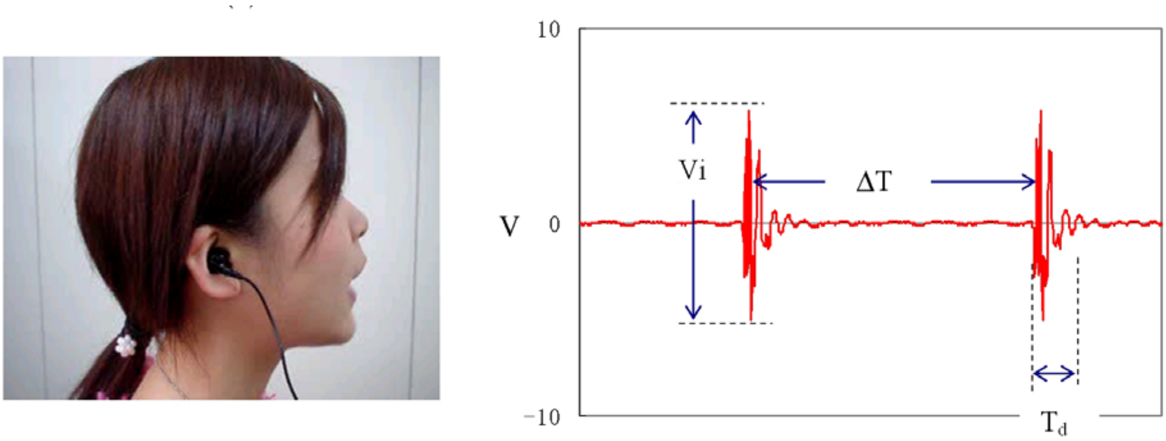
We used a bone conduction microphone to collect the tooth-touch sound signal. The bone conduction microphone cords used a highly sensitive vibration sensor, to collect vibrations that reached the skull from the site of the tooth-touch and convert this vibratory motion into audio signals. These devices could be worn all over the head to pick up vibrations, provided they were mounted on a relatively solid section of the head. Two types of bone conduction microphone could be utilized: an "ear microphone" that picks up vibrations in the auditory canal or a headset that captures vibrations on the head. In this research, we adopted the ear microphone. The bone conduction microphone had several qualities, including excellent noise tolerance (not affected by external noise) and flexibility when fitting. Figure 1(a) shows the way in which the ear microphone was fitted. In our experiment, we utilized a two-way communication device, EM7B-06, as the bone conduction microphone, developed by Temco Japan Co. The sound signal resulting from the tooth-touch was collected by the bone conduction microphone and amplified prior to low pass filtering and AD conversion. The signal was sampled at 10 [KHz] and processed for realization of the ECS. Figure 1(b) shows an example of the tooth-touch sound signal.

2.2 Characterization of tooth-touch sound signal

The characteristics of the tooth-touch sound signal varied between people. It was necessary to understand these characteristics in order to develop a man-machine interface using tooth-touch sound. In this section, we try to characterize the tooth-touch signal measured above. The characterization parameters shown in Figure 1(b) are defined as follows:

- Vi : Amplitude of tooth-touch sound
- Fc : Central frequency of tooth-touch sound using Fast Fourier Transform.
- T : The period between tooth-touch sound signals.
- T_d : Signal duration.

We collected the tooth-touch sound data from 17 adult volunteers aged 20 years old (9 males and 8 females). We asked them to generate 5 successive tooth-touch sounds in as consistent an interval as possible. Figure 1(b) shows an example of the tooth-sound signal. Table 1 summarizes the analyzed results of the above-mentioned characteristic parameters.



(a) *(b)*

Figure 1. *(a) Fitting the bone conduction microphone. (b) Example of measured tooth-touch sound signal.*

 ISBN 07 049 98 65 3

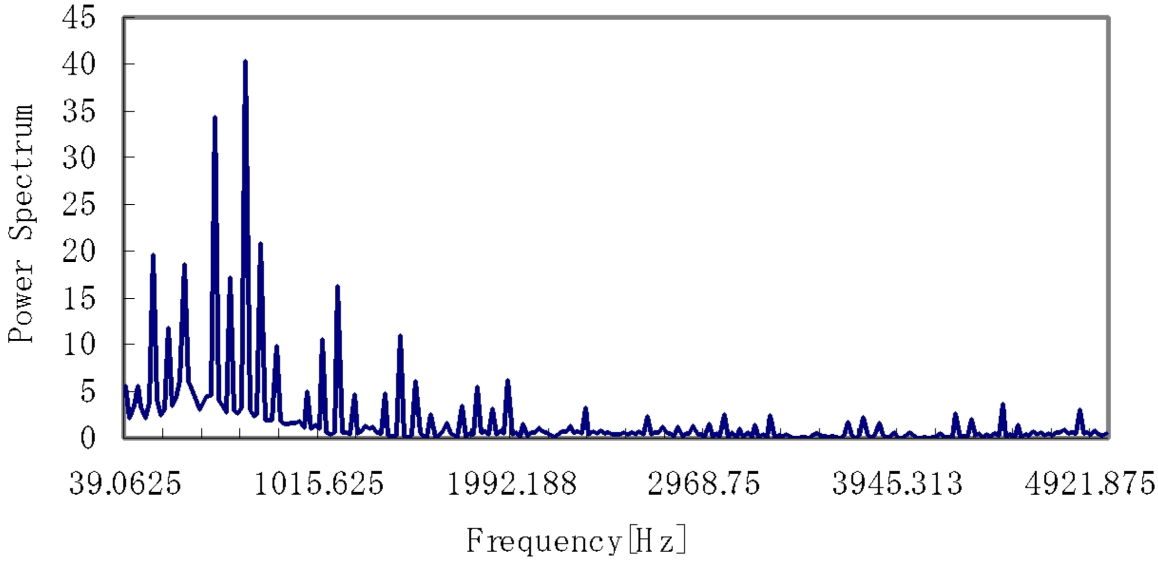


Figure 2. *Frequency spectrum of tooth-touch sound by FFT.*

The center frequency of the tooth-touch sound was about 600 [Hz] and its individual variation was very small. The duration of signals was less than 5 [msec]. It was, therefore, easy to process the signals and apply the data in controlling the ECS. Deviation in the amplitude and interval time of the 5 successive tooth-touch sounds was relatively small. In the next section, we outline how to learn about the mechanism involved in individuals' tooth-touch sound characteristics.

Table 1. *Resulting measurements of tooth-sound characteristics.*

Central Frequency Fc	Signal duration T_d	Deviation of tooth-sound amplitude Vi	Deviation of interval time between the tooth-touch sounds, ΔT
600 Hz	< 5msec	< 30%	< 20%

3. DEVICE ARCHITECTURE

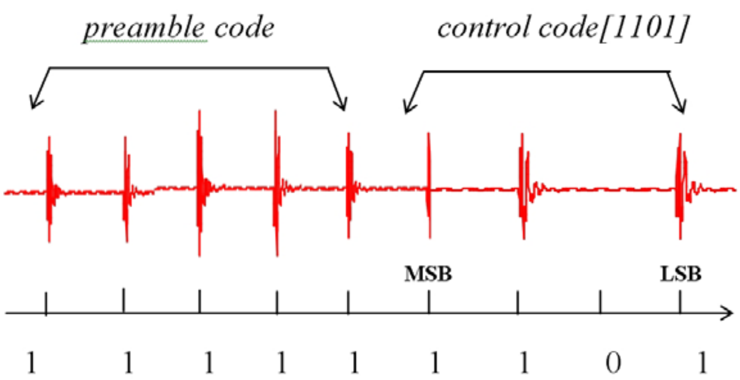
In this section, we propose a learning method for individuals' characteristic tooth-touch sounds. We then present a voice elimination algorithm and results from its simulation. Finally, we realize a "hands-free man-machine interface using the tooth-touch sound" and an FPGA chip.

3.1 Learning the characteristics of the tooth-touch sound

The characteristics of the tooth-touch sound have individual variations and vary with time, as shown in Section 2. To manage such variation, a device must learn and memorize changes. We proposed the learning algorithm as shown in Figure 3.5. Successive tooth-touch sounds, called "preamble code," were added prior to inputting the control code. In this example, the control code was [1101], which contained one start-bit "1" (the MSB of the control code) and one stop-bit "1" (the LSB of the control code). The mark, when the signal existed, was represented by "1". The space, when the signal did not exist, was represented by "0". Using the preamble code, the average amplitude (Va) and average period between signals (Ta) were calculated and stored in the memory, called the tooth-touch sound database unit. Setting the proper threshold levels for Va and Ta enabled us to distinguish the mark signal "1" and the space signal "0" from the tooth-touch sound signal. This learning process should be executed every time a user begins using the interface device and be updated, to adapt to changes in the user's physical conditions.

3.2 Voice elimination algorithm

Several kinds of noise, such as voice and ham noise, interfered with tooth-touch sound detection, the most serious being voice noise. The bone conduction microphone picked up not only the tooth-touch sound, but also the user's voice. Development of the voice elimination method was required to eliminate faults originating from background noise. Figure 4 depicts the signal containing both the voice and tooth-touch sound. The tooth-touch sound and voice were rarely generated at the same time. The tooth touch sound clearly resembled an impulse signal, having higher frequency components comparing with the voice signal and a distinct pattern. The voice eliminating method involved calculating the average of the absolute value of the signal. The average time was about 10 times longer than the duration of the tooth touch sound (from section 2 we set the time to 50 [msec]). Next, a threshold level to detect only tooth-touch signals had to be set. If the signal was larger than the threshold level, it could be regarded as the tooth-touch sound. The threshold level was set at a value ten times higher than the average of the voice signal.



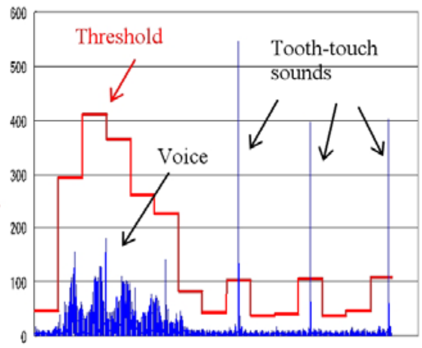


Figure 3. *Composition of preamble and control codes.*

Figure 4. *Voice elimination.*

3.3 Realization of man-machine interface using an FPGA chip

Figure 5 shows the FPGA based man-machine interface, with the above mentioned functions. The FPGA chip was operated at 3v-D.C. The tooth-touch sound detector and voice elimination unit eliminated noise from the digitized signal, leaving only the tooth-touch sound signal. The function of the learning unit was to learn and store characteristic parameters of individual's tooth-touch sound signals in the tooth-touch sound database. From the succession of tooth-touch sound signals in the database, the control codes for the ECS were generated and outputted from the chip.

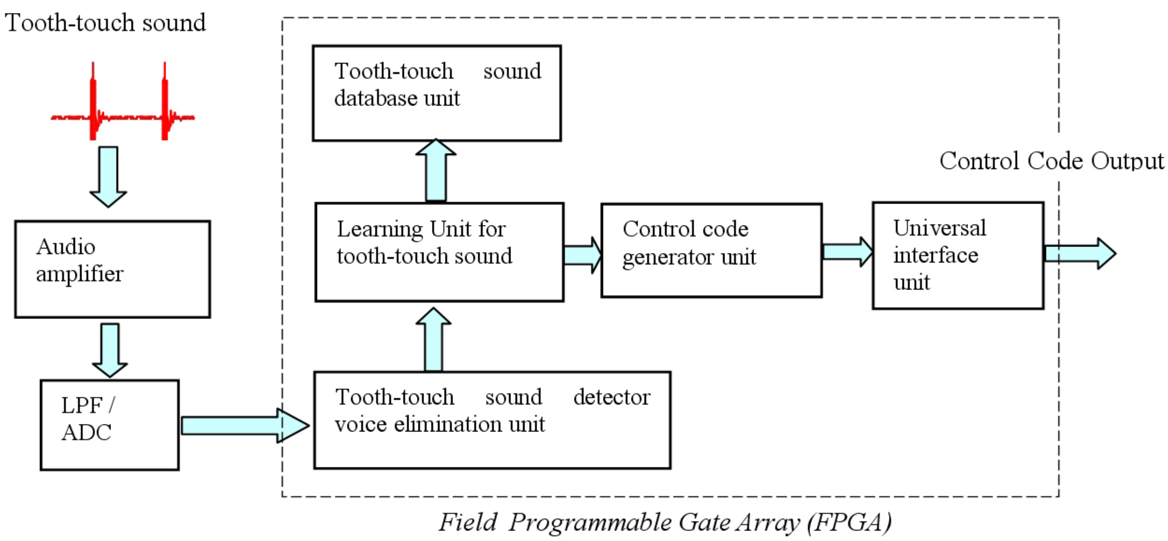


Figure 5. *Schematic diagram of an interface using tooth-touch sound.*

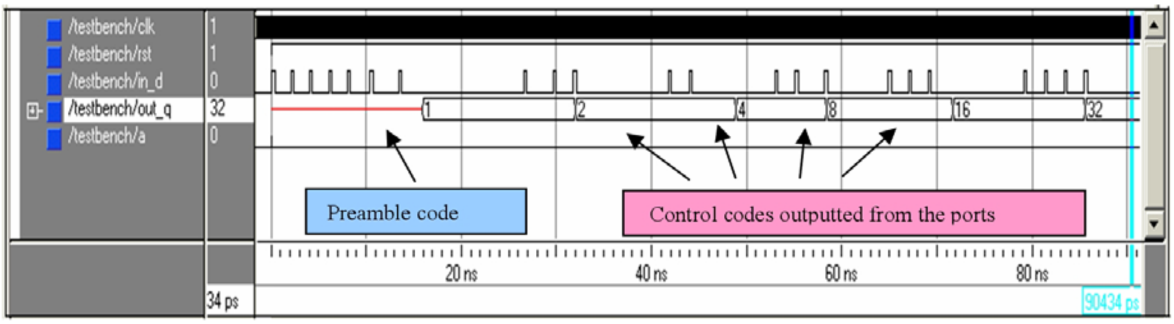


Figure 6. *Simulation result.*

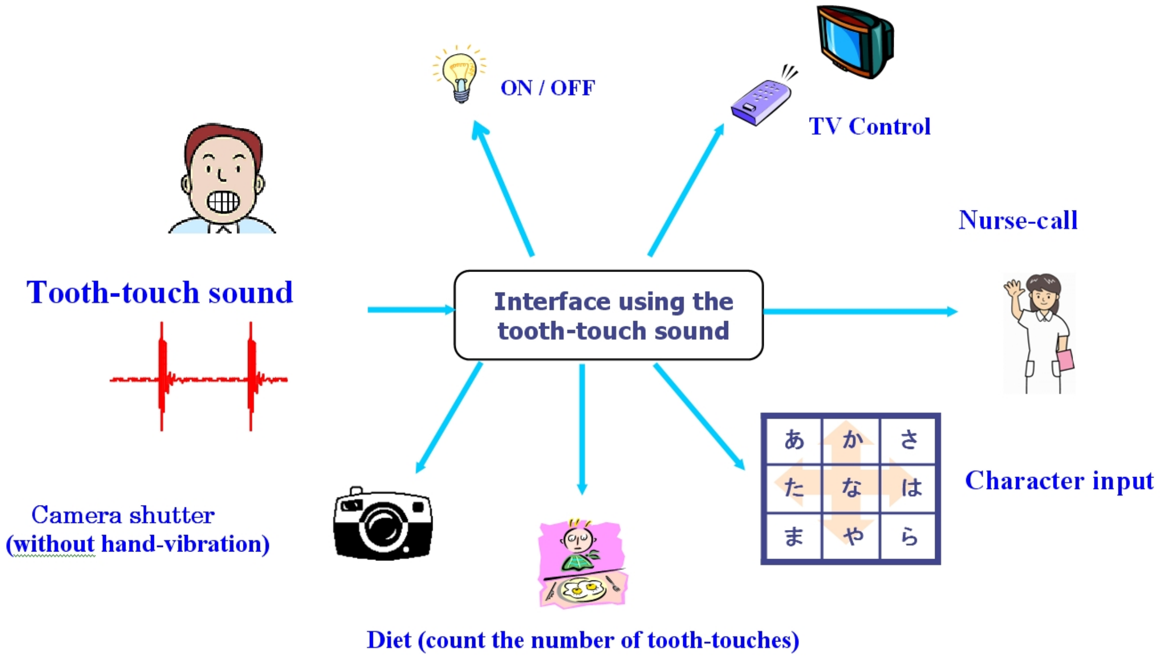


Figure 7. *Various applications of an interface using tooth-touch sound.*

3.4 Simulation results

To confirm the effectiveness of our system, we executed a timing simulation using logic simulation software, ModelSim. Figure 6 depicts the results of this timing simulation. The preamble code (5-successive pulses) and the following control code can be observed. The system could operate properly.

4. APPLICATIONS

A hand-free type, man-machine interface using tooth-touch sound, has several applications, as depicted in Figure 7. One obvious application is in an ECS, used to control consumer electronic appliances or to call a nurse in a hospital. Another type of application is as a character input device for communication aids.

In this research, we demonstrated a simple example of an ECS in which the interface was used to control an electric fan. We chose an electric fan because it could easily be controlled with only a few codes. Figure 8 shows an example having a 4-output control code system. Prior to inputting the control codes, users generated a preamble code consisting of the 5-successive tooth-touch sounds. This was to enable the device to learn the individual's characteristic tooth-touch sounds.

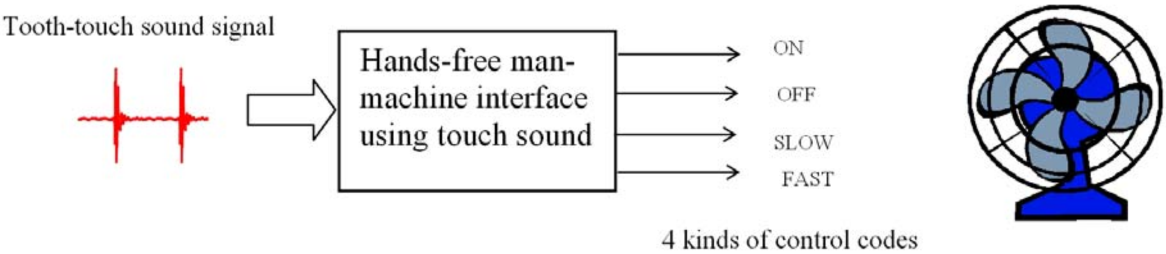


Figure 8. *Example of Simple ECS.*

5. CONCLUSIONS AND POTENTIAL IMPROVEMENTS

In this paper, we proposed a "hands-free man-machine interface device using tooth-touch sound". This device, which consisted of a bone conduction microphone, an audio amplifier, and an FPGA chip, had several advantages over existing technologies, including low cost and ease of handling. We analyzed the characteristics of the tooth-touch sound and showed that the tooth-touch sound was suitable for application in an ECS as a control signal. We then developed a learning method to allow for individual variations and a noise suppression algorithm. Parameters defining the amplitude and average time period between tooth-touch sound signals were used in the learning method. We tested our circuits using VHDL and developed a working FPGA chip. The prototype device was evaluated as an ECS, with the following shortcomings awaiting resolution: (1) insufficient suppression of noise caused by body movement, (2) extension of the device to other applications, (3) detailed evaluation of its usefulness for disabled persons.

6. REFERENCES

P. MacIsaac, A. Craing et. al. (2002), The mind switch environmental control system: Remote hands free control for the severely disabled, Technology and Disability 14, IOS Press, pp. 15-20.

D. J. Welling and J. Unsworth (1997), Fortnightly review: Environmental control systems for people with a disability: an update, British Medical Journal 315, pp. 409-412.

M. P. Barnes (1994), Switching devices and independence of disabled people, British Medical Journal 309 pp. 1181-1182.

A. Crag, P. MaIsaac et.al. (1997), Mind over matter: Brain signals and control of electrical devices, Today's Life Science 9, pp. 12-15.

L. Kirkup, A. Searle, A. Craig et. al. (1997), EEG system for rapid on-off switching without prior learning, Medical and Biological Engineering and Computing 35 pp. 504-509.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

Tongue-computer interface for disabled people

L N S Andreasen Struijk

Center for Sensory Motor Interaction, Aalborg University, DK-9220 Aalborg, DENMARK

naja@miba.auc.dk

www.smi.hst.aau.dk/~naja/

ABSTRACT

This work describes a new inductive tongue-computer interface to be used by disabled people for environmental control. The new method demands little effort from the user, provides a basis for an invisible man machine interface, and has potential to allow a large number of commands to be facilitated. The inductive tongue–computer interface implemented with 9 sensors was tested in three healthy subjects and the results shows typing rates up to 30 to 57 characters pr. minute after 3 hours of training.

1. INTRODUCTION

Tongue-computer interfaces are favorable since they are practically invisible and they are often manageable for people with even severe motor disabilities. A comparative study, comparing a tongue control method to a head control system and a rather simple mouth stick, resulted in all four severely disabled test persons preferring the tongue based control system, even though, it was not the fastest system (Lau and O´Leary, 1993), emphasizing the importance of aesthetics – the last thing many disabled persons want, is to be even more different. There have been different attempts to interface the tongue, including electrical contacts (Clayton et al, 1992), hall element techniques (Buchhold, 1995) and pressure sensors. Further, a current commercially available tongue control systems is based on pressure sensitive buttons placed in the mouth cavity over the tongue (New Abilities Systems, 1993). The use of electrical contacts may not function during eating and talking. The technique with the Hall element has similar limitations. Further, the use of pressure sensitive sensors does not seem optimal, since normal speech and swallowing generates tongue-palatal pressures in the range of 20-60% of maximal achievable pressure (Müller et al, 1984, Hayashi et al, 2002), which poses demands on the detection threshold and therefore may increase the risk of fatigue. The use of pressure-based sensors may limit the maximal number of sensors that can be placed in the oral cavity, since the requirement of pressure increases the tongue-palatal contact area. In addition, having only 9 control buttons the commercially available tongue control system (New Abilities Systems, 1993) far from utilizes the high selectivity in the movement of the tongue, which readily can pick out every single of our 32 teeth. Utilization of this selectivity would make, a variety of electric aids, including wheelchairs and neural prostheses, controllable with a wide range of commands from the same interface, making the tongue-computer interface suitable for environmental control systems and virtual reality systems.

Therefore this work describes a new inductive tongue-computer interface (ITCI) to facilitate tongue-activated commands.

2. METHODS

2.1 Theory

The detection method used in this work is based on Faraday´s law of induction for a coil, and uses variable inductance techniques.

From Faradays law the voltage drop across an inductance can be found as:

$$\varepsilon = -L\frac{di}{dt} = -\mu_0 \mu_r N^2 \frac{A}{l}\frac{di}{dt}$$

where

L = inductance

μ_0 = vacuum permeability

μ_r = relative magnetic permeability of the core material

N = number of turns

A = the area, and

l = is the average length of the magnetic flux path

When only air is present as the core of the inductance, μ_r=1. As a ferromagnetic material is placed in the coil, the core becomes a combination of air and ferromagnetic material, and μ_r changes according to the magnetic permeability of the ferromagnetic material.

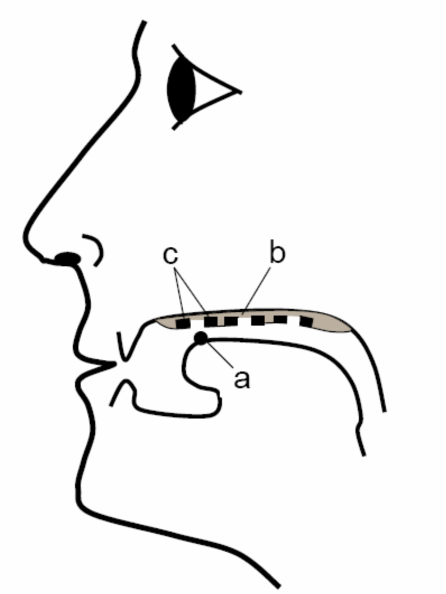


Figure 1. *The Inductive tongue control system. a: The activation unit, b: the palatal plate, c: the inductors. The tongue activates the sensors by placing the tongue-mounted activation unit at or inside a coil.*

Applying a sine wave current, i, of constant peak-peak amplitude, a constant amplitude voltage drop, ε, is obtained across the coil L. Introduction of the ferromagnetic material into the air gap of the coil, results in an increase of ε, which stays increased, a until the material is removed. This will be utilized for activation of a command in the inductive tongue control system (Fig. 1). The method resembles the known techniques used for displacement sensors (Göpel et al, 1989).

2.2 Mounting of Sensors at the Palatal Plate, and fabrication of the activation unit

Using dental acrylic, nine air cored inductors were placed on a palatal plate resembling the ones used as dental retainers, see Fig.2. The inner diameter of the inductor coils was 4mm, and the coils had 90-150 turns. A silicone tube was fixed to the palatal plate and carried the wires out of the mouth, see Fig. 2.

A small ferromagnetic metal cylinder was fabricated from stainless steel of the type SUS447J1. The diameter of the cylinder was 3.2 mm and the height was 2 mm. The steel had a maximum magnetic permeability, μr, of 2000-6500.

Figure 2. *Left: The palatal plate (placed on a mould of the upper part of the mouth) with inductive coils and a silicone tube leading the wires out of the mouth. Right: The activation unit glued to the tongue.*

2.3. Experimental Setup

The ITCS was tested in 3 healthy male subjects, age 22 -29 years.

The palatal plate with the inductors was placed at the hard palate. The activation unit was glued to the tongue using tissue glue, Fig. 2. A 50 kHz sine wave current with a 0.03 mA amplitude was applied to the inductors from a galvanically isolated current source. The signals form the 9 inductors were amplified and rectified. Then the signals from 4 of the inductors were inverted before the inductor outputs were connected to obtain 4 channels with 2-3 coils in series.

The subject activated the inductive sensors by positioning the tongue in a manner that placed the activating unit in the centre of the different inductor coils. hour on three consecutive days.

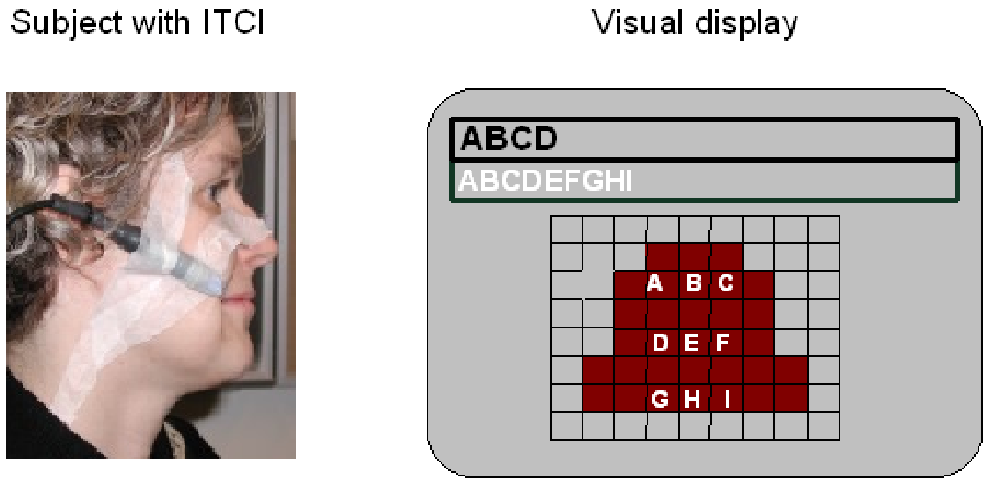


Figure 3. *Experimental set-up. The inductor leads comes out of the subjects mouth. The subject is provided with a visual feedback showing the position of the coils/characters on the dental palate, what to type in white and what have been typed by the subject in black.*

2.4 Signal Processing

The measured signals were amplified and rectified to obtain envelopes of the signals. Measurements were performed for intervals of 30 seconds

Then the signals were sampled and processed using the Matlab DAQ toolbox. The baselines of the sensor signals corresponding to no activation of the inductive sensors were subtracted. Using online thresholding of the inductor signals, one of the characters: "ABCDEFGHI" was related to the activation of each inductor, and typed on the visual display of a computer (Fig. 3) when the corresponding inductor was activated. From the visual display, the subject could see which sensor had been activated, see Fig. 3

In this way the subject was typing given sequences of the characters "ABCDEFGHI".

3. RESULTS

The subjects could activate desired sensors as shown in Fig. 4.

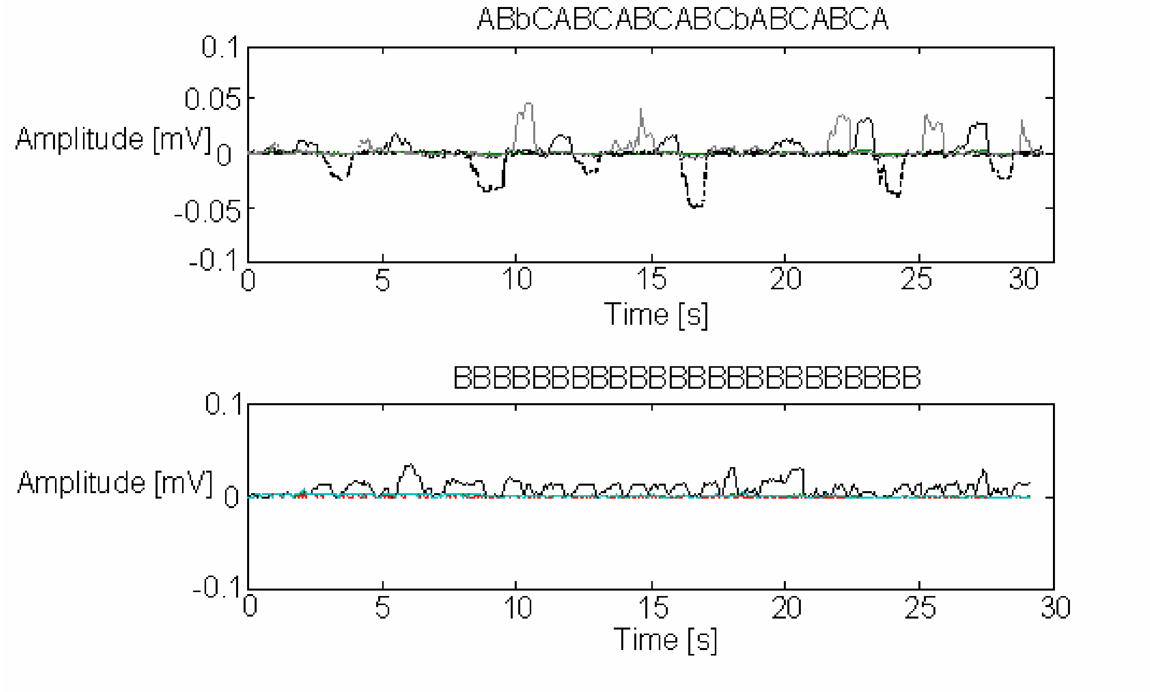


Figure 4. *Example of the signals from the ITCS during typing. The typed characters are shown above each of the two graphs. Capital characters are correctly typed characters, lower case characters are incorrectly typed characters. On the top graph, ABC is typed repeatedly:* ***Grey****: signal from the inductor related to the character "A",* ***black:*** *signal from the inductor related to the character "B",* ***dotted:*** *signal from the inductor related to the character "C". On the bottom graphs the signal from the ITCS is shown during repetitive typing of B. On both graphs signals from all 9 inductors are included, but only the activated inductors show amplitudes deviating from zero. The baseline potential corresponding to no activation of the coils has been subtracted. The data are from the third and final day of measurement.*

Typing sequences, including all nine characters, that had been repeated approximately 40 times during the 3 days of measurements, without having a visual display of the position of the characters resulted in a typing speed of 32-42 characters per minute with an error rate of 14-25%. Typing a random sequence of characters, that has not previously been typed, again without visual display, resulted in a typing speed of 30-57 characters per minute with an error rate of 15-29%. Repetitive typing of the same character was performed with maximal typing speeds of 48-114 characters per minute, and an error rate of 0-0.1%.

Table 1. *Typing rate of correct characters on day 3.*

Subject	correct characters pr. minute		
	max	mean	minimum
1	48	18.6	10.4
2	114	41	29.6
3	100	13.8	0

When asked on a scale from 1 to 10, about the perception of usage of the visual display (1= no use, 10=constant use) the subjects rated the use of the display to 1-3, the tongue palatal pressure needed for activation of the units (1=no pressure at all, 10=maximum obtainable pressure) to 2-3 and the difficulty in using the system to 3-5 (1=very easy 10 = impossible).

The typing rate of correct characters is shown in table 1. In average the mean typing rate of correct characters of all measurements from one day, improved with 117% from day one to day 2 and 22% from day 2 to day 3.

The total size of the designed inductors, given as the outer diameter was 5-6mm, indicating the possibility of having more than 25 sensors in the palatal plate of future ITCSs.

4. DISCUSSION

For a future control system to be truly successful it has, in reality to be a help for the user. This may imply that the system:

- Can be used/worn all day and night
- Is easy to use and induces a low degree of fatigue
- Is cosmetically acceptable in and outside the home of the user – preferably invisible
- Can be used to control a wide range of equipment, e.g. computers, wheelchairs, toys and prosthesis.
- Provides an efficient and quick activation of the desired function

These requirements may be met by future applications of this new inductive tongue-computer interface.

The small size of the sensor-coils opens up for the possibility of having the whole alphabet as separate "buttons" on the palatal plate, which may lead to substantial increase in e.g. the rate of writing for quadriplegics.

The typing rate of 30 to 57 characters per minute with 15-30% error rate suggests that the system may be quite efficient after sufficient training. The increase in the typing rate between consecutive measurement days indicate that learning may still taking place after the 3rd day, which may mean that longer training will reduce the error rate and further increase the typing rate. Therefore longer training is needed to get a more acceptable error rate.

The high typing rate of up to 114 activations pr. minute related to repetitive typing suggest the implementation of an interface having multiple functions for each inductive sensor, which can be activated by repetitive inductor activation.

The subjective experience of the subjects using the system suggests that the system may be used without visual display, potentially making it quite mobile. In addition the subjective experience was, that the activation of the inductive sensors only demands a low degree of tongue-palatal pressure, thereby reducing the risk of fatigue.

Future work will focus on implementing inductive tongue control interfaces with more than 9 inductive sensors and explore the possibility of implementation of a mouse or joystick function. Further, the Inductive tongue computer interface will be tested by people with motor disabilities. Finally, incorporation of wireless control and development of command strategies to control a wide range of devices will be considered.

Acknowledgements: We thank Professor Peter Svensson, for kindly being supportive with dental materials, Professor Osamu Okuno for support regarding the SUS stainless steel used for the activation unit and Helle I Mortensen for support with the experiments and the technicians at SMI for technical support. This work was made possible by grants from the Danish Technical Research Council, and Det Obelske familiefond.

5. REFERENCES

C Lau and S O´Leary, "Comparison of Computer Interface Devices for Persons with Severe Physical Disabilities", The American Journal of Occupational Theraphy, vol. 47, pp.1022-1030, 1993

C Clayton, R G S Platts, M Steinberg and J R Hennequin, "Palatal tongue controller", J. of Microcomputer Applications, 15, pp. 9-12, 1992

Buchhold N, "Apparatus for controlling peripheral devices through tongue movement, and method of processing control signals" ,United States Patent no. 5,460,186, 1995.

New Abilities Systems Inc, Patent: WO9307726, 1993.

E Müller, et al. ”Perioral tissue mechanics during speech production”. In J. Eisenfeld & C. DeLisi (Eds.) “Mathematics and computers in biomedical applications”. Elsevier Science publishers B.V., 1984”

R Hayashi, K Tsuga, R Hosokawa, et al., “A novel handy probe for tongue pressure measurement”, International Journal Of Prosthodontics 15 (4): 385-388 JUL-AUG 2002

W Göpel, J Hesse, J N Zemel, “Magnetic sensors”, in “Sensors, a comprehensive Survey” Volume 5,VCH, Verlagsgesellschaft mbH, D-6940 Weinheim, 1989.

ICDVRAT 2006

Session V. Rehabilitation and Route Learning

Chair: Charles van der Mast

Combining interactive multimedia and virtual reality to rehabilitate agency in schizophrenia

E A Lallart, S C Machefaux and R Jouvent

Emotion Center CNRS Paris 6 UMR 7593
Hôpital de la Salpêtrière, 47, Bd de l'Hôpital 75013, Paris, FRANCE

eliselallart@yahoo.com, sebastienmachefaux@yahoo.fr, rjouvent@ext.jussieu.fr

www.cnrs.fr

ABSTRACT

New interactive technologies offer the opportunity to involve the user's body in a virtual environment while seeing herself/himself performing the actions. Interactive exercises with a video-capture reinforce the perception-action loop, which is the pillar of agency (i.e. the ability to attribute the intention of an action to its proper author). We present a new paradigm as a possible treatment of agency disturbances in schizophrenia.

1. INTRODUCTION

New information and communication technologies play an increasingly important role in society. These innovations confront humans with the necessity of adapting cerebral function by interacting with a non-human intelligence. Likewise Virtual Reality (VR) offers the opportunity to realize an experiment in an environment which is possible to control by presenting limited sensory information from the real world or by replacing some information with a virtual world. Within the past decade Virtual Reality has been widely used for rehabilitation of certain pathologies and has shown surprising efficacy. The benefits of VR is applied to treat patients who suffer from a stroke (Jack, 2001), anxiety (North, 1998) and autism (Strickland, 1997). Our research laboratory began recently to apply these new methods to schizophrenia as well (Jouvent & Rautureau, 2005).

The schizophrenic dissociation comprises many symptoms which have in common a disintegration, psychic disorganization and depersonalization. For instance patients who suffer from schizophrenia may experience a loss of their self, meaning that their self is an artificial entity to them and that their mental experiences are disconnected. According to several authors schizophrenia might be a "pathology of consciousness", since individuals with schizophrenia have alterations of consciousness in specific cognitive abilities, such as executive functions.

Thus the delusion of influence is characterized by the fact that the individual does not feel in control of her/his movements and thinks that she/he is a passive instrument in the hands of an external will. Another example of schizophrenic dissociation is the delusion of reference in which patients may believe that they are controlling others' movements.

Those symptoms of "extraneity" have in common the fact that schizophrenia may include a deficit in the consciousness of action and in the attribution of their or other's actions. The ability to correctly attribute the intention of an action to its proper agent is called "sense of agency". Several authors state that the disruptive mechanism of agency is one of the essential etiopathogenic mechanisms of schizophrenia (Frith, 1992). Joëlle Proust suggests that what is defective in schizophrenia is probably not the rational thinking, but rather the self-attribution of intentions.

Only a very few studies were administered using VR with schizophrenia patients (Jeonghun, et al. 2003). However we believe that a combination of VR and multimedia technologies could be an efficient tool to rehabilitate a certain aspect of the disorder: agency. Objections could be made, stating that VR could present a potential risk for patients as it mixes reality to an artificial world, and therefore it may erase the boundaries between reality and a virtual world. However no reliable evidence supporting this view has been made so far (Da Costa, 2004). Instead this risk may be inverted into a potential cue for rehabilitation of self –perception in the environment.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

To support our hypothesis we will first present a summary of the theoretical background on schizophrenia and define the concept of agency. Then we will present a validation of a study which measures the sensory-motor performances of a group of controls in two virtual reality tests. Finally we will open up new horizons for a potential rehabilitation in schizophrenia using our paradigm.

2. THEORETICAL BACKGROUND

A certain view of the neurophysiology is that the ability to attribute an action, an idea or an intention, depends on the control of agency. In the literature, two distinct theories mainly articulate this idea: the theory of shared representations and the theory of self-monitoring.

2.1 Theory of shared representations: recognition of action

In social contexts it is difficult to attribute an action to its proper author. This statement is supported by the neurological evidence of the mirror neurons discovered in the monkey by Rizzolatti (1996). Those premotor neurons (F5) are activated both when the monkey observes some one acts and when he is the author of the action. For Marc Jeannerod this experiment is an illustration of what he names the “shared representations”. What allows one to attribute actions to its proper author is the “Who system”, which distinguishes auto generated actions from others (Georgieff, Jeannerod, 1998). Jeannerod hypothesizes that in schizophrenia this system is deficient and leads to symptoms such as the syndrome of influence.

2.2 Theory of self-monitoring

Beside the distinction between the first and third person in action, a more physiological model proposed by Christopher Frith explains how one differentiates the sensory consequences of willed actions from the sensory information provided by the external world. The individual uses his motor commands according to the “internal efferent model” (Ito, 1970; Wolpert, 1995): when a motor instruction is sent to the muscles, a copy of this instruction – the efference copy (Von Holst, 1954) – is also sent to a comparator or a self monitoring system. Held 1961, suggests that the efference copy is sent to the comparator where it is stored and compared to the reafferented information (for instance proprioceptive or visual) on what sort of movement was made. According to Campbell, the basis of agency is the matching at the level of the comparator between the efference copy and the sensory feedback of the movement. In other words what gives the feeling that one is the author of a movement is when the efference copy has received the instruction to move the person’s arm and matches the movement that one perceives.

If the feedback does not correspond to the expectation of the individual, this latter corrects his action by means of the perception-action loop until the outcome of the movement corresponds to the desired movement.

In schizophrenia, several studies showed the difficulty for schizophrenic patients to monitor an action when they are deprived of an actual sensory feedback (due to a deficit of the self monitoring) (Frith & Done, 1989). Experimental studies involving a distortion of the feedback showed a deficiency for schizophrenic patients to detect a difference between their willed action and the actual sensory feedback, causing symptoms such as the syndrome of influence. According to C. Frith, there is a breakdown in the mechanism of the efference copy and in the comparator which provokes a dysfunction of the sense of agency. Schizophrenic patients seem unable to monitor their motor instructions (Frith, Blakemore, Wolpert 2000). They rely on the visual feedback instead of the efference copy of the motor commands to predict the success-result of their actions. Proust states that they apply a control called “error-control”, instead of applying a “cause-control” (Proust, 2003). In this theory the self attribution of action is essentially assured by the forward motor model, which predicts the motor and sensory consequences of the action. This model of control of action can be directly applied to the problem of action recognition. Action recognition can rely on the concordance between a desired action and its sensory consequences.

If patients with schizophrenia perform a test repetitively in which they are agents, and perceive themselves being an actor, adjusting their actions to match the feedback they perceive (perception-action loop), it might help restructuring the integrity of their body in action. That is the reason why we propose a task in which the subject has a visual perception of her/his body performing the action as if he were facing a mirror.

2.3 Hypothesis

In this paper we hypothesize that the exercises we propose are sensitive enough to observe a difference when we alter the image of the user. If this is the case, it would underline the importance for the subject to use a visual perception of his body in action to increase the performances.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

3. METHOD – VALIDATION IN A CONTROL GROUP

3.1 Participants

We recruited a total of twenty two controls in this study who were evaluated by the Mini International Neuropsychiatric Interview (Lecrubier et al. 1997). Participants were nine males and thirteen females ranging in age from 18 to 35 years old, with an average age of 23,87. The exclusion criteria were the following: familial antecedents of schizophrenia and bipolar disorder, neurological history, stroke.

3.2 Material

We use Augmented Reality tests: the subject's image and his surrounding are projected by a web cam in a virtual world; therefore there is an association between the real and the virtual world by using the integration of real images (RI) with virtual entities (VE). The user interacts with this virtual world by means of his movements. In Eye Toy, the RI and VE are displayed simultaneously on the same screen (latency less than 100 ms).

One of the most important characteristics in using augmented reality is the way it makes a transformation of the locus of interaction possible. The interactive system is no more understood as a face-to-screen exchange, but dissolves itself in the surrounding space and objects. Those tests are on line with our problematic on agency because the user is not only an agent, but also sees himself being the actor. Two virtual environments were used in this system:

Wishi Washi: the user faces a virtual window which is covert by soap. The task is to wash each window by the means of the movements of the arms and body as fast as possible. Each time a window is washed, a new one is presented to the subject. The total time for each test is two minutes (Fig. 1).

Figure 1. *A subject performing the Wishi Washi task.*

Figure 2. *A subject performing the Mirror test (with horizontal and vertical inversion).*

In the Mirror test the subject is facing a virtual mirror which contains a bonus or a malus in each corner. Sometimes the mirror is horizontally or vertically inverted. Each subject has to catch the bonus and disposes of three minutes for each test. If the subject touches more than three malus, the test ends straight away (Fig 2).

3.3 Procedure

All subjects start with the Wishi Washi test − which does not involve any distortions of the visual feedback − and then do the Mirror test − which includes distortions. Three practice trials were given at the beginning of both tests to train subjects to the sessions.

In each test we alternated between two conditions: a control condition and a second one in which we hide the central part of the body by means of a rectangular object placed in front of the projector, so that the subject only sees his arms moving but has no visual feedback of the other parts of his body. We call this condition the "mask" condition.

Therefore half of the subjects did the following sequence:

- 3 tests of familiarisation
- 2 control tests – 2 tests with alteration of the body image (“mask” condition) – 2 control tests – 2 “mask” tests
- The other half:
- 3 tests of familiarisation
- 2 “mask” tests – 2 control tests – 2 “mask” test – 2 control tests

The performances of each test were recorded and analysed using SPSS. In Wishi Washi, the score represents the number of windows washed and the bonuses obtained; in Mirror it is the number of balls caught. In both tests the dependant variable is the total score obtained.

4. RESULTS

In both tests and in each condition, the distribution of the total score among the subjects is globally Gaussian, as shown by the following graphics:

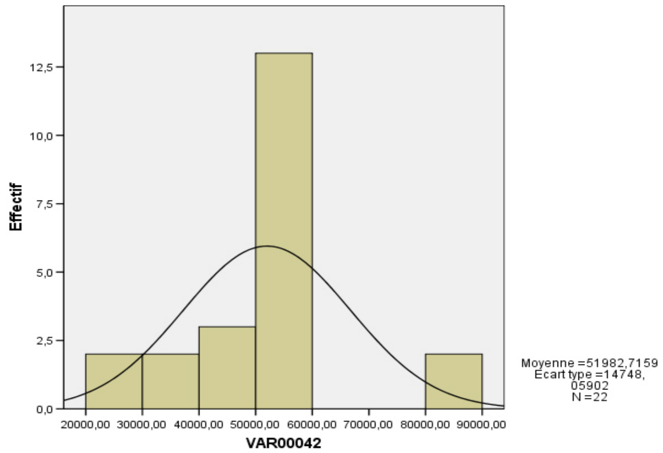


Figure 3. “*Wishi Washi*”:
Distribution of the averages of the total score. Control condition.

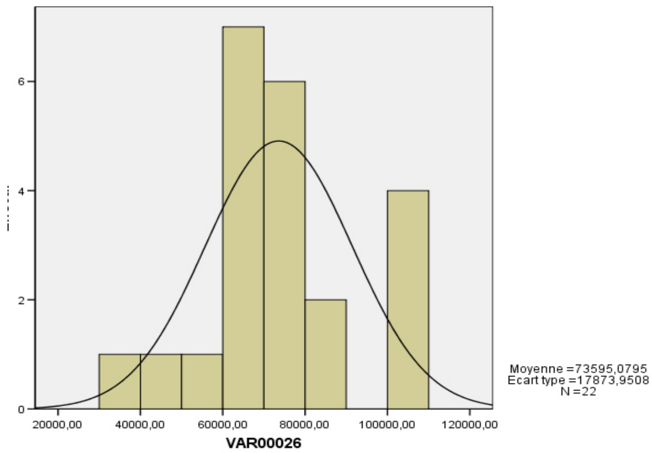


Figure 4. “*Wishi Washi*”:
Distribution of the averages of the total score. “Mask” condition.

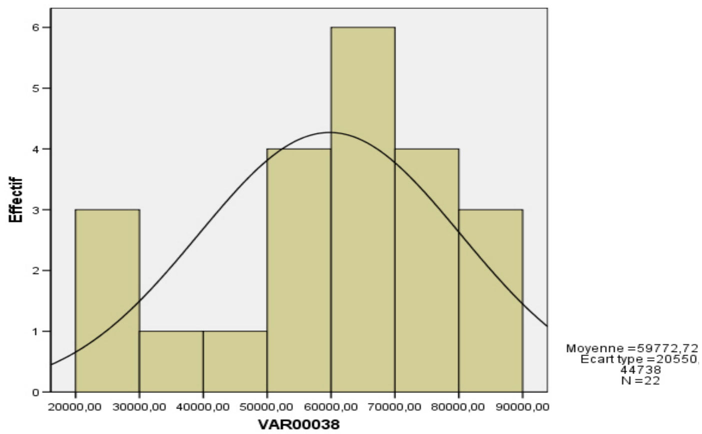


Figure 5. *“Mirror”*:
Distribution of the averages of the score. Control condition.

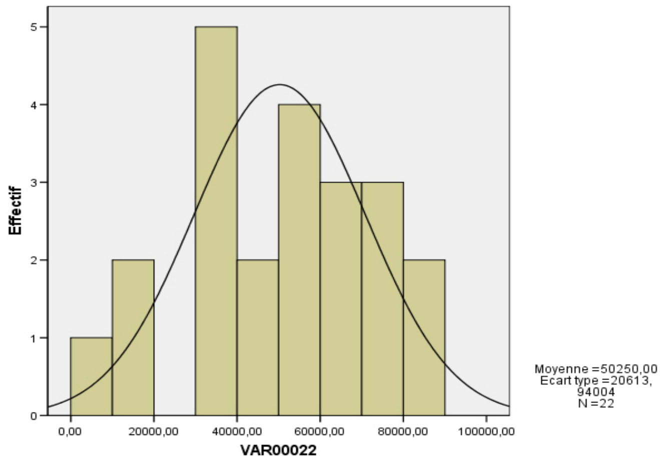


Figure 6. *“Mirror”*:
Distribution of the averages of the score. “Mask” condition.

Results show that the presence of the “mask” provokes significant changes in the performances; subjects do better in the Wishi-Washi ($p < 0.0001$) test and worse in Mirror ($p < 0.0001$).

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

Table 1. *Wishi Washi.*

Conditions	Score	SD	t-test	p
Control	13 635	384.76	35.44	<0.0001
Mask	17 875	384.76	46.46	<0.0001

Table 2. *Mirror.*

Conditions	Score	SD	t-test	p
Control	62 186	2669.08	23.30	<0.0001
Mask	51 077	2669.08	19.14	<0.0001

5. DISCUSSION AND PERSPECTIVES

5.1 Different levels of action

Our work is a validation of a paradigm to use VR and multimedia technologies in schizophrenia patients as a way to rehabilitate the self in action. In this present paper, we gathered the scores in a population of controls in order to further compare them to schizophrenia in a further study. In both tests the twenty two controls presented homogeneous performances. Each test corresponds to a different level of action: one requires sensory-motor ability (Wishi-Washi) whereas Mirror corresponds to a higher level of action accompanied with a better control of the body movements.

According to the literature, there are three distinctive levels of action: the first one is sensory-motor, the second is contextual; the last level of action is the episodic memory (Koechlin, 2003). According to several studies, the contextual condition is defective in schizophrenia but not the sensory-motor stage (Green, 2005). We therefore believe that schizophrenia subjects would have performances in the average range in the Wishi Washi test, but be particularly impaired in the other test. Furthermore they would do worse in Mirror with the mask, whereas the performances would be the same with the mask.

5.2 Body ownership and agency

In those tests the matching of visual and proprioceptive signals contributes to the inter-modal sensory image of the body. Experimental data show the prevalence of vision over the other senses in self recognition: we feel our hand where we see it, and not the opposite. The sense of the position is recalibrated in order to be conformed to the visual information (especially in the Mirror test, with inversion of the image).

Botvinick and Cohen (1998) placed a plastic hand in front of subjects, which was hidden by a screen. A tactile simulation was applied simultaneously to the real hand and to the plastic one. After a while subjects had the illusion that they felt touched whereas it was the plastic hand – which they were seeing. In other words, the subject had a sense of ownership of the plastic hand just by seeing it.

Therefore watching ourselves performing an action, like in the tests presented, may contribute to our sense of body ownership. If one sees an image in front of him/her which moves when she/he moves and that the two movements are congruent, then the image seen must own to him/her. The position sense is recalibrated to conform to the visual information. The visual feedback of the body in action is also a way for individuals who experience a syndrome of influence to feel that they are the cause of the action (feeling of agency), since there is a correlation between the expected actions and the result (Frith, 1992).

5.3 Seeing oneself in action

The purpose of hiding the central part of the body (the mask condition) is to make a switch between the first and third person perspective as it changes the visual presence of the self. Indeed with the "mask" the user needs to make abstraction of a part of himself. In this condition, the integrity of the body is parcelled out and the subject does not adjust his actions to the perception of his body in action any more, but only to his arms, isolated from his body.

On the one hand, in the test without distortion (Wishi-Washi), subjects don't need a higher cognitive ability to perform it because it is a simple motor test. The "mask" does not disrupt the performance in this task.

On the other hand the Mirror task -with a distortion of the visual feedback- requires a higher cognitive ability and various mental rotation levels. Indeed it is necessary for the subject to have a mental representation of his body in space. The "mask" presents a real handicap because it alters the inhibition ability of the user to accomplish the task. We believe that schizophrenia patients could benefit from seeing themselves in action in order to participate in the restructuring of their self and body-in-space representation.

In our study we have presented that the combination of virtual and multimedia technologies are applicable to a group of controls and have a noticeable sensitivity. We have tested two schizophrenia subjects using our paradigm to measure the feasibility of this study. We intend to apply this paradigm to a group of schizophrenia patients.

6. CONCLUSION

The finality of Virtual Reality combined with multimedia technologies is to allow sensory-motor activities with a visual feedback of the user's body in action. The process of body perception and the creation of one's own body based on multiple stimuli (proprioceptive and visual) is a prerequisite for rapid and effective action with our surroundings and an awareness of being the agent. The objective of the present study was to validate a paradigm for agency rehabilitation in schizophrenia. In the future we would like to demonstrate that repetitive practice of the models presented can help restructuring the perception-action loop and thus give access to a rehabilitation of the feeling of agency in schizophrenia.

7. REFERENCES

D Jack, R Boian, AS Merians, M Tremaine, GC Burdea, SV Adamovich, M Recce, H Poizner (2001), *Virtual reality-enhanced stroke rehabilitation*, IEEE Trans Neural Syst Rehabil Eng. 9(3):308-18.

MM North, SM North, JR Coble (1998), *Virtual reality therapy: an effective treatment for phobias*, Stud Health Technol Inform, 58:112-9.

D Strickland (1997), *Virtual reality for the treatment of autism*, Stud Health Technol Inform, 44:81-6.

R Jouvent, P Canet, G Rautureau (2005), *Which reality for which psychopathology?* In: VRIC – Laval Virtual (Richir S, Taravel B, eds), pp 121-124. Laval.

C Frith (1992), *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, USA : LEA Publishers.

K Jeonghun, et al. (2003), *A Virtual Environment for Investigating Schizophrenic Patients' Characteristics: Assessment of Cognitive and Navigation Ability*, CyberPsychology & Behavior, Vol. 6, No. 4: 397-404.

RM da Costa, LA de Carvalho (2004), *The acceptance of virtual reality devices for cognitive rehabilitation: a report of positive results with schizophrenia*, Comput Methods Programs Biomed, 73(3):173-82.

G Rizzolatti, Fadiga, L., Gallese, V. and Fogassi, L. (1996) *Premotor cortex and the recognition of motor actions*, Cog. Brain Res., 3: 131-141.

N Georgieff, M Jeannerod (1998), *Beyond consciousness of external reality: a "who" system for consciousness of action and self-consciousness*, Conscious Cogn, 7(3):465-77.

M Ito (1970), *Neurophysiological aspects of the cerebellar motor control syste,* Int. J. Neurol. 7:162-176.

DM Wolpert, Z Ghahramani, MI Jordan (1995), *An internal model for sensorimotor integration*, Science, 269, 1880-1882.

E Von Holst (1954), *Relations between the central nervous system and the peripheral organs,* Br. J. Animal Behavior, 2 89-94

R Held (1961), *Exposure-history as a factor in maintaining stability of perception and coordination,* Journal of Nervous and Mental Diseases, 132: 26-32.

C Frith, D Done (1989), *Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action,* Psychol Med, 19(2):359-63.

C Frith, S Blakemore, D Wolpert (2000), *Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action*, Brain Research Reviews, 31:357-363.

J Proust (2003), *Thinking of oneself as the same*, Conscious Cogn., 12(4):495-509.

E Koechlin, F Kouneiher (2003), *The architecture of cognitive control in the human prefrontal cortex*, Science, 302(5648):1133.

M Greene, P Uhlhaas., M Coltheart (2005), *Processing and social cognition in schizophrenia*, Current psychiatry reviews, 1, 11-22.

M Botvinick, J Cohen (1998), *Rubber hands "feel" touch that eyes see*, Nature, 391:756.

Investigating the efficacy of a virtual mirror box in treating phantom limb pain in a sample of chronic sufferers

C D Murray[1], E Patchick[1], S Pettifer[2], T Howard[2] and F Caillette[2]

[1]School of Psychological Sciences/[2]School of Computer Sciences,
University of Manchester, Oxford Road, Manchester, UNITED KINGDOM

craig.murray-2@manchester.ac.uk, *emma.patchick@manchester.ac.uk*

[1]*www.psych-sci.manchester.ac.uk/staff/CraigMurray*, [2]*aig.cs.man.ac.uk*

ABSTRACT

This paper describes a pilot clinical study to evaluate the efficacy of using immersive virtual reality (IVR) as a rehabilitative technology for phantom limb pain experienced by amputees. This work builds upon prior research which has used simple devices such as the mirror box (where the amputee sees a mirror image of their remaining anatomical limb in the phenomenal space of their amputated limb) to induce vivid sensations of movement originating from the muscles and joints of their phantom limb and to relieve pain. The IVR system transposes movements of amputees' anatomical limbs into movements of a virtual limb which is presented in the phenomenal space of their phantom limb. The primary focus here is on a qualitative analysis of interview data with each participant throughout the study. We argue that the findings of this work make a case for proof of principle for this approach for phantom pain treatment.

1. INTRODUCTION

Phantom limb pain (PLP) is the chronic experience of pain in the residual impression of a limb which persists following amputation. Research has shown incidence rates as high as 85% with PLP sufficiently severe to require withdrawal from social or work environments for considerable periods of time (Sherman et al., 1984). The relationship between phantom limb pain (PLP), prosthesis use and psychological well-being is an intimate one. For instance, significant correlations have been observed between adjustment to amputation and pain (Katz, 1992), with adjustment to amputation less likely as levels of pain increase. Amputees with PLP are less likely to use a prosthetic limb (Dolezal et al., 1998). Non-prosthesis use often results in the restriction of normal activities (such as self-care, visiting friends and carrying out domestic work), and is associated with higher levels of depression (Williamson et al., 1994; Murray, 2005).

Phantom limb pain has thus far proved to respond poorly to the current range of pharmaceutical, surgical and psychological treatments. However, one promising development in this regard was Ramachandran's mirror-box (Ramachandran and Rogers-Ramachandran, 1996) which allowed amputees to view a reflection of their anatomical limb in the visual space occupied by their phantom limb. Ramachandran reported evidence that the box induced vivid sensations of movements originating from patients' phantom limbs and in some cases relieve their PLP. The mirror box has also been used with similar success in lower-limb amputees, where viewing a reflection of an anatomical limb in the phenomenal space of a phantom limb resulted in a reduction in phantom pain (MacLachlan, MacDonald and Waloch, 2004).

The mirror box presents a number of inherent limitations in treating PLP, highlighted previously by the authors (Murray et al., 2005). The illusion is tentative, relying on the patient maintaining attentional focus on the reflected image as opposed to the moving anatomical limb. If the patient were to move away from the mirror, or not to keep the mirror within visual range, the illusion would be compromised if not completely broken. It also operates within a narrow spatial dimension, relying on the patient remaining in a restricted, fixed position. Finally, the possible movements that can be induced in the phantom limb are constrained by the need for patients to imagine themselves carrying out two-handed tasks which are concordant with synchronous mirror-image movement; conducting with both hands, for example. Mirror-box work sometimes uses two-handed tasks so that the patient can focus on both limbs (intact and reflected). With single handed-tasks it becomes more difficult for the patient to ignore the visual information coming from their intact limb.

However, the mirror box work is promising and suggests that other visual therapies that work in similar ways, while overcoming some of its limitations, may also relieve PLP. In the present project we have implemented a similar phenomenon to the mirror-box using immersive virtual reality (IVR) (Murray et al., 2005; Murray et al., 2006a). This system transposes the movements made by their intact anatomical limb into movements of a virtual limb in the phenomenal space occupied by their phantom limb. Unlike the mirror-box, IVR allows the patient to perform tasks without having to remain visually and spatially fixed with a relatively narrow dimension. The tasks implemented in IVR can be more complex, and require a wider range of anatomical movements, than possible with the mirror-box. A further advantage is the ability of an IVR system to implement single-handed tasks, and there is the potential to implement tasks similar to those used in normal physical rehabilitation contexts. Therefore, the IVR system allows complex fine and gross motor movements and thereby the possibility for users to engage in tasks made impossible by the mirror box. In an earlier paper we presented initial findings regarding three patients (Murray et al., 2006b). In the remainder of this paper we outline the experiences of 5 patients regarding changes in the phenomenology of their phantom limb or phantom pain. The objectives of the present work are to show proof of principle for the development of this kind of technology for the treatment of phantom limb pain and to stimulate further work in this area.

2. METHODOLOGY

2.1 Design

Due to the relatively nascent approach of using IVR to treat PLP, the research is a pilot study to assess proof of principle for such systems. Participants used the system on a weekly basis, with the precise intervals between sessions determined by participant availability. This made it difficult to standardise the procedure precisely: however all participants used the system at least 7 times, with a maximum of 10 sessions. Each IVR session lasted 30 minutes. Participants completed four tasks in repetitions: placing the virtual representation of the phantom limb onto coloured tiles which light up in sequence; batting or kicking a virtual ball; tracking the motion of a moving virtual stimulus; and directing a virtual stimulus towards a target. It was hypothesised that participants would experience phenomenological changes in their experience of their phantom limb which would lead to significant short-term and long-term reductions in the frequency and severity of PLP. The reduction of phantom limb pain was hypothesised to lead to significant positive changes in psychosocial issues, activity restriction, and satisfaction with a prosthetic limb.

2.2 Participants

Participant characteristics are shown in Table 1.

Table 1. *Characteristics of Study Participants (N=5).*

Name	*Gender*	*Age*	*Site of Amputation*	*Time Since Amputation*	*Duration Suffering PLP*	*Prosthesis Wearer?*
PK	Male	63	Left Arm Above Elbow	12yrs 10 months	12 yrs, 10 months	No
WW	Male	60	Right Leg Below Knee	12 yrs 3 months	12 yrs	Yes
BH	Male	56	Left Arm Below Elbow	39 yrs 8 months	39 yrs, 8 months	Yes
SM	Female	61	Right Leg Above Knee	11 yrs 8 months	11 yrs, 8 months	No
DT	Female	65	Left Arm Below Elbow	1yr 3 months	1 yr	Yes

Participants were recruited through Manchester's sub-regional Disablement Services Centre and contacted on the basis of having phantom limb pain and being adults without any major visual or cognitive deficits. For upper-limb amputees only left-arm amputees could be included (the equipment has a right-handed glove). Both left and right lower-limb amputees could be recruited. Participants were a minimum 12 months post-amputation to ensure the phantom pain experience was chronic in nature. Eight participants were recruited after the inclusion criteria were exercised. Of these, three withdrew from the study after a maximum of 3 sessions. One participant was advised to withdraw participation by his physiotherapist due to weakness in the remaining anatomical active limb. The remaining two participants withdrew due to difficulties arising in transportation to and from the research site. All participants had undergone varied and extensive previous treatments for their phantom limb pain; in the case of PK, this included the implantation of a deep brain stimulator which had subsequently malfunctioned. Interestingly, a majority of participants (3 out of 5) had also been treated using the mirror box with no success. As such, the final sample comprised of participants suffering with particularly stubborn phantom limb pain.

2.3 Materials

A V6 virtual reality head-mounted display was used to present the computer-generated environments to participants and to facilitate immersion. In order to monitor and represent participants' arm/hand/fingers or leg/foot movements a 5DT-14 data glove and sensors were used for upper-limb amputees, while sensors were used for lower-limb amputees. Sensors attach to the elbow and wrist joints or the knee and ankle joints for upper- and lower-limb amputees respectively. A Polhemus Fastrak monitors head movements and arm/leg movements. A virtual environment represents the participant within a room, from an embodied point-of-view (see Figure 1). A full virtual body representation is provided for participants. A virtual representation of the phantom limb is made available by transposing the participant's opposite anatomical limb (e.g. their physical right arm) into the phenomenal space of their phantom limb (e.g. their virtual left arm – see Figure 2).

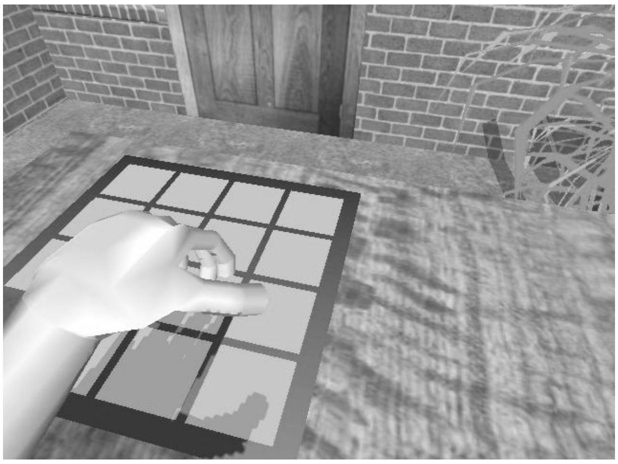


Figure 1. *One possible view participants may see when taking part in the experiment.*

Figure 2. *A person with a left-arm amputation taking part in the study.*

Transferring movement from an anatomical limb into movement of a virtual limb on the opposite side of the body is possible due to joint angles parameterization. For example, once the joint angles are recovered from the right physical arm, applying these joints angles to the left virtual arm results in mirroring the movement. This method of transferral as well as other implemented software generates responsive, fluent, real-time motion, allowing virtual limbs to move in synchrony with anatomical limbs. The representation of the virtual body gives realistic results at a gross level, which the authors and participants found to be satisfactory for successful immersion. However, in an attempt to reduce discrepancy, the interface on start-up does allow the colour of skin and clothes to be altered to approximate those of the participant (see Murray et al., 2005, 2006a for more details).

2.4 Data Collection

The small sample size in the study reported here precludes a meaningful quantitative analysis. Therefore in the present paper we emphasize the importance of achieving a qualitative understanding of patient's phantom limb experience, and of their experience of using the IVR system, in their own words. The broad approach here is a phenomenological one; with the aim of understanding patients' own embodied experiences (see Murray, 2004). This is achieved through semi-structured interviews carried out at each session (lasting about 15 minutes each). Alongside this qualitative data the below quantitative information was also collected and is reported where appropriate in the results to contextualize the qualitative findings.

Measures of PLP were monitored over the course of the study period through a number of questionnaires: The McGill Pain Questionnaire (MPQ) (Melzack, 1975) which was completed on the first visit and a follow-up 2 weeks after completion of participation in the project; The Short-Form MPQ which was administered following each interim session; and Pain Diaries were also completed daily throughout the course of the intervention to allow a contextualised analysis of participant's phantom pain experience. The Trinity Amputation and Prosthetic Experience Scales (TAPES) (Gallagher and MacLachlan, 2000) is a multidimensional self-report instrument designed to help understand adjustment to an artificial limb and was administered on the first visit and at the 2-week follow-up. Finally, the Amputee Body Image Scale (ABIS) (Breakey, 1997) assesses how an amputee perceives and feels about his or her body experience and was administered at the same time interval as the TAPES. These additional measures were completed in order to assess any secondary benefits that may have occurred alongside pain reduction.

3. RESULTS

3.1 DT

DT had no volitional control over her phantom arm and her phantom pain was localized to the phantom hand. She described the pain in her fingers as, "almost as if they're trapped. My fingers are trapped. The knuckles hurt." The phantom hand remained paralysed in a fist: "The palm of my hand hurts, I think because it is as if they're [fingers] sticking into it." Her phantom pain was "there all the time" and often interrupted her sleep and everyday life. Her initial scores on the MPQ, using weighted rank values were 37.25 (out of a possible 88.21). DT reported vivid sensations of phantom limb movement during the tasks and, after just one session, that her index phantom finger had been somewhat freed: "It's funny… one of my fingers is coming out." It is unlikely that DTs self-reports were confabulation since she expressed extreme surprise at this new experience, "It's weird… I know that it can't be and your brain's telling you it is and you know damn well that the brain's wrong and it can't be! The brain's weird!" After the fourth session the phantom pain was reported to have eased overall: "the pain has gone down a bit and the [pain] flashes have gone down a bit… so it's been quite good for the last few days."

One negative aspect associated with DT's participation was that she usually experienced slightly more intense phantom pain for a period immediately after sessions. When asked how long this would tend to last, DT stated, "about 24 hours. Yes, it spiked a lot and goes up but then it improves slowly". This verbal report was confirmed through observation of DT's pain diaries. Despite the negative aspects of DT's experience, the trend of improvement continued over the study period. By the final session both the phantom thumb and index finger had been released and DT reported having some voluntary control over these digits. DT surmised that this had a positive affect on the phantom pain in general: "I think the fact that it's bought my fingers out – at least some of them out – has helped the pain considerably. It feels more comfortable." DT did not complete follow up measures; in lieu of the full MPQ follow-up evaluation, it is interesting to note that the Short Form MPQ, showed an overall evaluative decrease in phantom pain following the first and last sessions from 'Discomforting' to 'Mild'.

3.2 SM

SM suffered violent phantom pains on a regular basis which often left her: "passed the screaming stage… you end up crying." Her phantom pain at its worst was described as, "electric shocks" which travelled down the leg and across her foot. This sensation would build up until, "as one if firing off, another one's following." She had good voluntary control over the movement of her phantom leg and foot and found that sometimes movement would "interrupt the pain." Her phantom pain often interrupted her sleep and everyday life and initial scores on the MPQ were 77.84 (out of a possible 88.21). SM reported a lowering of PLP during the IVR sessions: "It kind of diverts the mind away from the pain," and a transferral of sensations into her phantom leg throughout the tasks: "I was moving the limb itself and trying to get into the position you would actually use it – you know, to kick the ball." She enjoyed the experience of using the system: "the right leg was trying to do it for me. I think it's a good exercise". However, her PLP increased following sessions for a period of up to 48 hours, which she attributed to the "stimulation of phantom nerves".

Whilst SM enjoyed using the system and "exercising" her leg, in general, both qualitative and quantitative reports of her levels of PLP did not seem to alter much throughout the period of the study. SM's follow up questionnaires were not returned so comparative MPQ scores are not available. However, the Short-Form MPQ which was completed at each visit showed an overall evaluative decrease in phantom pain from 'Distressing' at the first session, to 'Discomforting' at the last session.

3.3 PK

PK suffered severe phantom pain "twenty-four seven. I'm never, ever out of pain" in 12 years and 3 months since amputation. He experienced his pain as a "burning, cutting pain – like someone cutting me with a hot knife." He also had a vivid experience of a strap around his wrist that was "pulled really tight" and his hand was paralysed in a cup position with the fingers always being very painful. Initial scores on the MPQ were 46.41 with ABIS scores at 33 (out of a possible 95). (TAPES Scores are not reported here as PK did not wear a prosthetic limb and therefore could not complete this questionnaire.) PK had little to no volitional control over movement in his phantom limb and could only swing it side to side with movement of the stump.

PK reported vivid sensations of a transferral of kinaesthetic sensations into his phantom limb whilst using the equipment which "allowed me to forget that my [phantom] arm was actually in a fixed position." After the 3rd session, the use of the system felt "more like reality than virtual reality". PK also reported that, whilst using the system, "it took away a lot of my phantom pain." For the first few sessions, PK – like DT and SM - reported increased levels of phantom pain after sessions. However, PK attributed this to the pain returning after a lull which increased subjective experience of the pain: "having had nothing *(during sessions)* and then having the pain, it feels stronger."

After 4 weeks of using the system, PK was very surprised to report his phantom limb moving of its own accord for a period of 1 hour whilst he was at home. During this time, he was "virtually pain free." It is unlikely that this report is due to confabulation since PK telephoned the experimenter whilst this was occurring to share his extreme surprise. Towards the end of his involvement PK reported, with some surprise, a number of changes in the phenomenal experience of his phantom limb which improved his phantom pain. The strap around his wrist had loosened: "before, the strap was so tight that my fingers felt swollen up and really, really painful. Now that strap seems to be not as tight, it feels as if I've got circulation." He could make very small volitional movements of his phantom fingers and had some control over the orientation of his wrist. Finally, he reported the experience of telescoping in his phantom arm which had a beneficial effect on the pain: "My limb actually seems shorter… I don't know why, it just seems to be shrinking." PKs follow-up MPQ scores fell to 32.76, with the ABIS scores also decreasing to 25.

3.4 WW

WW has suffered with severe PLP for 12 years and 10 months and experienced intense pain in the sole of his phantom foot on a regular basis, "as if someone's ramming a knife in". WW also reported experiencing many different kinds of pain in the phantom foot which he could attribute to previous pain experience in the right foot before amputation including: a broken ankle; a burn to the top of the foot, and even the memory of his toes being tightly squeezed when he was a child due to small shoes, amongst other things. WW's initial MPQ score was 44.54. The ABIS evaluation showed a body image disturbance score of 57. The TAPES measures of psychosocial adjustment, activity restriction and prosthesis satisfaction were 42, 20 and 40.

WW suffered with simulator sickness which meant the first session was terminated early. At one point during the second session, his anatomical left leg collided with his stationary prosthetic leg. WW commented that this was an "uneasy sensation… it looks on the thing [HMD] like it's not in the way but then you bang into it and it feels queer." Asked to try and elaborate on this, WW mentioned his PLP had increased slightly during this period. This is consistent with research which sites sensory-motor incongruence as a possible source for painful sensations (McCabe et al., 2005). WW chose to remove his prosthesis during subsequent sessions to avoid this situation and consequently, he engaged more and reported decreased feelings of nausea.

WW reported vivid sensations of movement in his phantom leg during sessions: “It’s a queer sensation…I’m doing the games with my right leg” and expressed pleasure at feeling as though he was “achieving” something with his phantom limb. After 3 sessions, WW used his experience of the IVR system to begin self-hypnosis; a technique which he had once previously used to aid pain control. He would be “imagining myself on this machine and it seems to help a bit that I can look down and see my leg.” It seemed the virtual representation of the body helped WW to focus his concentration. He began self-hypnosis sessions 3 times a day using this technique; a factor which may confound the findings of this research.

WW reported that the use of IVR system was helping his PLP: “the burning pain is abating a little bit. So it’s improved a little bit.” When reflecting on the experience at his final session, WW referred to an easing in the various different types of pain he experienced: “it seems to have taken the edge off them. You know, they’re not as severe.” However, quantitative measures at follow up would suggest that the PLP has worsened rather than improved. The follow-up MPQ score increased to 61.79, with ABIS increasing to 62. The TAPES measures of psychosocial adjustment, activity restriction and prosthesis satisfaction remained largely the same at 45, 20 and 40 respectively. It is difficult to consolidate WW’s verbal reports with these findings. WW’s pain did appear to be the most inconsistent of all participants since he did not suffer with just one type of phantom limb pain; the pain memories would come and go at random intervals, making it difficult for him to comment on his pain over any extended period of time as it would fluctuate so greatly. This highlights a difficulty in comparing pain scores obtained at one point in a study with another point (when either of these particular data collection times may not be typical of the person’s pain level).

3.5 BH

BH self-reported that his phantom pain: “doesn’t bother me regularly”, but it was particularly stubborn in that it had persisted for almost 40 years. The pain he experienced was severe cramping in the phantom hand and his initial MPQ score was 21.72. The ABIS evaluation showed a body image disturbance score of 56. The TAPES measures of psychosocial adjustment, activity restriction and prosthesis satisfaction were 48, 18 and 39 respectively. BH had very good voluntary control over his phantom limb before use of the system but reported highly unexpected sensations in the experience of his physical body whilst using the equipment: “It felt like I was leaning over to do it.” He reported vivid sensations of feeling as though his right physical arm was leaning over his body towards the left hand side in order to make the virtual left, or phantom arm, move. This was a highly unexpected experience since his physical right arm would always remain on the right hand side of his body. It is highly unlikely that this self-report was due to confabulation since he continued to report this experience even after he knew that this was not possible: “ I feel like I’m reaching right over but I can’t possibly be because it’s like a mirror so my arm should physically be going that way!” Interestingly, he also had strong impression of his physical phantom arm moving whilst carrying out the tasks and found it difficult to accept the contrary: “it feels like I’m moving my left arm. But according to you, I’m not!”

In the last 3 weeks of involvement, BH reported no experience of PLP (previous reports and diaries had shown at least 2 episodes per week). Whilst BH could not conclusively attribute this improvement to use of the system, he did comment: “I’m not doing anything different from what I’ve always done… and I’ve not had the cramp since.” The follow up questionnaires showed improvements on all measured scales apart from the TAPES prosthesis satisfaction. MPQ scores dropped to 11.89 and ABIS to 46. The TAPES psychosocial adjustment subscale improved to 52, with a small decrease in perceived activity restriction to 11.

4. DISCUSSION

Analysis of the qualitative data does provide opportunities for tentative conclusions to be proposed that are beyond the scope of quantitative statistical tests. All participants made some reference to a transferral of sensations into their phantom limbs during testing. This is a particularly interesting finding when we consider the proportion of paralysed phantom limbs which could not be moved voluntarily. It may be, in fact, that this treatment would be most beneficial for those with paralysed phantom limbs as some phantom pain can be directly attributed to the inability to move paralysed phantoms into comfortable positions.

The reporting of sensations of movement in phantom limbs appeared to be more vivid for upper limb amputees. This finding could reflect the greater degree of movement afforded by the virtual hand and fingers as opposed to the virtual foot. The very nature of feet is less dexterous than hands so that authors feel this is a situation that is difficult to avoid in virtual reality systems. It would however, be possible to develop specific tasks using virtual lower limbs which encourage the user to manipulate the foot in a more detailed way. For example, tasks could be made more difficult to force participants to use their feet in more dextrous ways. It would also be interesting to use a virtual representation of a foot, as opposed to a shoe, which may make the lower virtual limbs more analogous to the upper virtual limbs and reduce any discrepancy between the experiences of lower-limb and upper-limb amputees when using the system.

DT reported the most drastic change in phenomenological experience of her phantom limb. After the first session, changes in the once paralysed phantom limb began to help relieve aspects of her phantom pain experience, as recorded in qualitative reports. A speculative hypothesis could explain this in terms of a greater plasticity in the brain for more recent amputees as it has had less time to re-define the internal model of the body and to cortically reorganise, which is strongly correlated with phantom limb pain (Flor et al, 1995). As such, it is possible that this system may be of more benefit for more recent amputees. However, PK also reported surprising changes in his experience of his phantom limb after over 12 years of paralysis so this could suggest that any potential improvements may be caused by some other factor than time since amputation. It would be interesting to couple this research with some form of function brain imaging technique to help establish if and how much normalisation is occurring through use of the system.

Another common finding between many participants was the experience of an increase in the level of phantom pain which followed sessions. This is difficult to consolidate with existing research on phantom limbs. As PK pointed out, it may be that the easing in pain during sessions, that almost all participants commented on, means any subsequent pain feels more severe. All pain experiences are relative and subjective and a constant level of pain may be easier to overcome than fluctuating levels of pain, as is the case when pain levels were unexpectedly lowered during use of the system. It could also be that the increased concentration required to carry out the novel tasks actually have some temporarily detrimental affect on absolute phantom pain. This is an issue that would be closely monitored in future research.

As mentioned above, all participants did make reference to a decrease in experienced phantom limb pain whilst immersed in the virtual environment. This is a positive result which should be investigated further. SM specifically used the word “distraction” when reporting this reduction, which leads us to the conclusion that the novelty of the tasks may provide an easy, temporary escape from the phantom pain. It is important to carry out further testing, not only with a larger sample size, but with a control condition in order to assess any placebo affects of pain reduction caused by the novelty of the task. A suitable control condition for this research would be the use of the IVR system without any transposition of movement in the virtual world i.e. physical right leg movements would correspond to virtual right leg movements (see Murray et al., 2006a).

There are also practical issues to consider with this research. Two participants withdrew from the study due to difficulties in travelling to research site. If this equipment were to be disseminated, it would require dedicated disablement service centres running regular, intensive trials with participants. This may present some difficulties for certain patients and is a factor to consider when planning dissemination if justified through further research. One participant dropped out due to pre-existing weakness in his remaining anatomical which meant he found using the IVR system difficult. Whilst the virtual tasks themselves were not judged to be particularly strenuous by this participant, the novelty of the tasks meant he was using his limbs in a way they were not used to. Obviously, it is crucial that use of this system does not exacerbate pain of any kind in any way. To ensure this is the case, it would be simple to introduce parameters which control the difficulty and nature of the tasks to accommodate users of varying ability. A larger database of tasks would also provide the flexibility required when treating amputees of all ages and levels of health.

Finally, a crucial factor to be addressed in future research would be the intensity of the intervention. Previous work with the mirror box has used regular intervention sessions of up to twice daily (MacLachlan et al, 2004; Ramachandran, and Rogers-Ramachandran, 1996). Using this protocol, participants came for sessions on a weekly basis which may be insufficient to facilitate change. This is understandable given that the majority of participants had suffered with phantom pain for over eleven years; it may be unrealistic to expect a weekly intervention for less than 3 months to have any dramatic effect on phantom pain. This is especially the case with this kind of intervention which is not only novel for participants to get used to but also novel in terms of how IVR has been used in rehabilitation in general. Two participants (DT and PK) expressed an interest in coming for testing more often as they felt that it would be of more benefit to them. Such decisions to decrease (or increase) intervals between interventions may impact on the practical issues of the general reliability and availability of participants and would require consideration of these factors.

Overall, the majority of participants report some kind of positive impact on their phantom limb pain throughout testing. The authors feel that the initial qualitative analysis coupled with the positive trends in the quantitative data show sufficient proof of principle to justify further research with the system. Future work would first require some adaptations to the system to accommodate the various changes specified above. Randomised controlled trials with a larger sample size would be crucial to assess the efficacy of this system in treating phantom limb pain, not only over and above any placebo effects but also in order to extrapolate to the wider amputee population.

5. REFERENCES

J.M. Dolezal, S.H. Vernick, N. Khan, D. Lutz and C. Tyndall (1998), Factors associated with use and nonuse of an AK prosthesis in a rural, southern, geriatric population, *International Journal of Rehabilitation and Health*, **4**, *4*, pp. 245-251

H. Flor, T. Elbert, S. Knecht, C. Wienbruch, C. Pantev, N. Birbaumer, W. Larbig, and E. Taub (1995), Phantom limb pain as a perceptual correlate of cortical reorganization following arm amputation, *Nature*, **375**, *8*, pp. 482-484.

J. Katz (1992), Psychophysiological contribution to phantom limbs. *Canadian Journal of Psychiatry*, **37**, pp. 282-298.

M. MacLachlan, D. McDonald, and J. Waloch (2004), Mirror treatment of lower limb phantom pain: a case study, *Disability and Rehabilitation*, **26**, pp. 901-4

C.S. McCabe, R.C Haigh, P. Halligan and D.R. Blake (2005), Simulating sensory-motor incongruence in health volunteers: implications for a cortical model of pain, *Rheumatology*, **44**, pp. 509-516

R. Melzack (1975), The McGill pain questionnaire: major properties and scoring methods, *Pain*, **1**, pp. 277-279.

C.D. Murray (2004), An interpretative phenomenological analysis of the embodiment of artificial limbs, *Disability and Rehabilitation*, **26**, *16*, pp. 963-973

C.D. Murray (2005), The social meanings of prosthesis use, *Journal of Health Psychology*, **10**, *3*, pp. 425-441.

C.D. Murray, S. Pettifer, F. Caillette, E. Patchick and T. Howard, T. (2005), Immersive virtual reality as a rehabilitative technology for phantom limb experience, *Proc. IWVR 2005: The 4th International Workshop on Virtual Rehabilitation*, California, pp.144-151.

C.D. Murray, E. Patchick, S. Pettifer, F. Caillette and T. Howard (2006a), Immersive virtual reality as a rehabilitative technology for phantom limb experience: A protocol. *CyberPsychology & Behavior*, **9**, *2*, pp.167-170.

C.D. Murray, E. Patchick, F. Caillette, T. Howard and S. Pettifer (2006b), Can immersive virtual reality reduce phantom limb pain? In *Medicine Meets Virtual Reality 14: Accelerating Change in Healthcare: Next Medical Toolkit* (J.D. Westwood, R.S. Haluck, H.M. Hoffman, G.T. Mogel, R. Phillips, R.A. Robb, & K.G. Vosburgh, Eds) IOS Press, pp. 407-412.

V.S. Ramachandran and D. Rogers-Ramachandran (1996), Synaesthesia in phantom limbs induced with mirrors. *Proceedings of the Royal Society of London, B Biological Sciences*, **263**, pp. 377-386.

R.A. Sherman, A., Sherman, C.J. and Parker, L. (1984), Chronic phantom and stump Pain among American veterans: results of a survey, *Pain*, **18**, pp. 83-89.

G.M. Williamson, R. Schulz, M.W. Bridges and A.M Behan (1994), Social and psychological factors in adjustment to limb amputation, *Journal of Social Behavior and Personality*, **9**, pp. 249-268.

Use of a virtual-reality town for examining route-memory, and techniques for its rehabilitation in people with acquired brain injury

J Lloyd[1], T E Powell, J Smith and N V Persaud[2]

Department of Psychology, Birmingham University, Edgbaston, Birmingham, UK

[1]*jxl340@bham.ac.uk, t.powell@bham.ac.uk, etc, jennie.smith@hmps.gsi.gov.uk,* [2]*natedawg2k@hotmail.com*

www.psychology.bham.ac.uk

ABSTRACT

Route learning difficulties are a common consequence of acquired brain injury, and virtual environments provide a novel tool for researching this area. A pilot study demonstrated the ecological validity of a non-immersive virtual town, showing performance therein to correlate well with real-world route learning performance. The first patient study found that a rehabilitation strategy known as 'errorless learning' is more effective than traditional 'trial-and-error' methods for route learning tasks. The second patient study, currently in progress, will assess whether naturalistic route learning strategies of map and landmark use can be combined effectively with errorless techniques. A final study will investigate the relationships between route learning performance and scores on a select battery of neuropsychometric tests.

1. INTRODUCTION

Impaired wayfinding ability is a little-investigated cognitive deficit associated with many cases of acquired brain injury (ABI). In one of the few studies to estimate the prevalence of wayfinding difficulties in ABI, Barrash et al (2001) found that almost a third of all patients sampled experienced new difficulties finding their way since their injury, and this figure rose to 86% for individuals with injury to particular cortical regions (including the hippocampal and parahippocampal area and parts of the medial and inferotemporal cortices). Wayfinding is an important skill for everyday life, and the question of how best to rehabilitate individuals who suffer wayfinding impairment merits investigation.

It is possible that practical difficulties arising in the study of route learning have contributed to a lack of research in this area. The problem of finding a route near to the recruitment site that is novel to all participants is a difficulty common to many studies of navigation (e.g. Kirasic et al, 1984), but may be particularly important where participants have specialised transport needs and may not be able to easily travel to distant sites. Physical impairments of some participants with ABI could also make learning long or outdoor routes difficult or fatiguing, and deficits in concentration or executive function could make experiments into route learning in the outside world potentially hazardous. However, in recent years, virtual reality has been used successfully in a number of route learning studies (see Darken and Peterson, 2001 for a review) and circumvents many of the practical difficulties associated with real-world studies (e.g. Haq et al, 2005).

Brooks (1999) successfully used a virtual simulation of a real-world scenario to help a patient with severe amnesia learn routes around a hospital by consistent repetition. Virtual reality is particularly suited to assessing the impact of multiple trials on memory, as it tends to allow for more repetitions of a route in a given time than a real-world setting (e.g. Darken and Banker, 1998), and also will not cause the physical fatigue possible in multiple repetitions of a real-world route.

The present series of experiments also employed a virtual-reality town to investigate memory for novel routes in people with acquired brain injury. The overall aim of the study was to compare the usefulness of various route-learning techniques, and in particular to compare errorless learning, a technique thought to stimulate implicit (or procedural) memory with trial-and-error memory, a technique relying upon more explicit strategies. While several previous studies have employed virtual reality environments to assess wayfinding, the majority have either used simple, maze-like environments (e.g. Grön et al (2000)), have focused on direct transfer of spatial learning from a virtual simulation of a real-world space to the real space

 ISBN 07 049 98 65 3

itself (e.g. Brooks et al, 2000), or have used virtual environments as a means for conducting brain imaging scans to assess the functional neuroanatomy of wayfinding (e.g. Hartley et al, 2003). The present study, by contrast, sought to use a complex and lifelike virtual environment as a stand-alone research tool, in which to compare route-learning strategies.

This article first describes the virtual environment used for the investigation of route learning. A study conducted to test the ecological validity of the virtual town is then reported. Three studies using the virtual town to investigate route learning techniques for people with ABI are then described. The first of these, which is completed, is a comparison of two techniques known as errorless and errorful learning. The second, which is in progress, is an investigation of the benefits of combining errorless learning with naturalistic strategies for route learning. The final study, also currently in progress, looks at the relationships between neuropsychometric test profiles and route learning performance under various conditions.

2. THE VIRTUAL ENVIRONMENT

A desktop, interactive computer-generated virtual town was employed. The software was called 'Driv3r', developed by REFLECTIONS Interactive Limited, an Atari Studio. Recent growth and investment in the computer gaming industry means that mainstream computer games offer some of the most convincing and sophisticated virtual environments available, along with being relatively inexpensive. The virtual town selected from Driv3r was based upon the real-world town of Nice, and contains a large network of streets and buildings. The game also features drivers and pedestrians moving around the town in realistic patterns.

3. EXPERIMENT 1: ECOLOGICAL VALIDITY OF THE VIRTUAL ENVIRONMENT

3.1 Introduction

While several studies have demonstrated good general equivalence between virtual and real world route learning (e.g. Koh et al, 1999; Skelton et al, 2000), virtual environments can vary considerably, and so a pilot study was conducted to assess the ecological validity of the specific virtual town selected for the present series of experiments. Participants' route learning in the virtual town was compared, using several measures, with their route learning in the real world.

3.2 Method

3.2.1 Apparatus. The software was played via a Sony 'Playstation 2 games console', connected to a 19-inch colour television screen via SCART cable. Participants were seated at a distance of approximately one metre from the screen. A standard, manually operated 'control-pad' was used by the experimenter to control movement through the town. Whilst individuals were told to imagine they were driving through the town, the experimenter actually moved the virtual character in a 'running' motion, down the centre of the road. A first-person viewpoint was selected so that the participant could not see the body of the character, and was free to imagine travelling in a car. It was desirable for participants to imagine being in a car because (a) this kept consistency with the real-world study in which people travelled by car, and (b) moving in the middle of the road gave a better view of relevant landmarks and junctions than moving along the pavement.

3.2.2 Procedure. 14 neurologically healthy individuals were 'driven' around a route within the virtual reality town described above, before being returned to the start point and asked to try and direct the experimenter, verbally, around the route they just saw. They were also taken by car to a residential area in the real-world city of Birmingham, UK, where they were driven around a route of similar length and with the same number of turns. They were then returned to the route start-point and asked to direct the experimenter around the route, in the same way as in the virtual task. The order in which the virtual and real-world tasks were administered was counterbalanced to control for order effects.

3.3 Results

A paired-samples t-test showed no significant difference between numbers of wrong turns taken on the real versus the virtual route ($t = -.563$, $df = 13$, $p = .583$). Furthermore, a strong correlation was observed between the number of errors made in the virtual and real towns ($r=.807$, $n=14$, $p<0.0005$). There were also statistically significant correlations ($p<0.05$) between the extents to which individuals reported using several popular strategies in the real and virtual environments. Individuals appeared to use cognitive maps, landmarks, a position-encoding process known as 'dead reckoning' and guess-work to similar extents in both settings.

3.4 *Conclusions*

These results suggest that the virtual task is a good approximation of real-world route learning, not only through demonstrating an impressive correlation between performance across real and virtual versions of the task, but also through illustrating individuals' tendency to use common wayfinding strategies to comparable extents across tasks. The ecological validity of the virtual town demonstrated in this study, coupled with support in the literature for the ecological validity of virtual environments in wayfinding studies in general (e.g. Koh et al, 1999) supported the decision to use the virtual environment in the following experiments.

4. EXPERIMENT 2: A COMPARISON OF ERRORLESS VERSUS ERRORFUL ROUTE LEARNING IN FOR PEOPLE WITH ACQUIRED BRAIN INJURY

4.1 *Introduction*

Errorless learning is the name given to learning episodes in which effort is made to prevent errors from occurring during the 'encoding' stage of information processing. It can be contrasted with the well-known 'trial-and-error', or 'errorful' learning method in which individuals improve their performance by learning from their mistakes. It is thought that errorless learning works by exclusively strengthening *correct responses* in memory, thereby minimizing competition or interference from any possible incorrect responses when one attempts to remember or use the information (i.e. at the 'retrieval' or 'response' stage).

Errorless learning has been used with particular success with individuals with acquired brain injuries (e.g. Baddeley and Wilson, 1994), as well as with people with other cognitive deficits, such as learning difficulties (Sidman and Stoddard, 1967) and dementia (e.g. Winter et al, 1999). It has been suggested that people with certain cognitive impairments benefit more from errorless techniques because where they are unable, or less able than neurologically healthy individuals, to consciously correct errors that occur during learning, errorless learning allows for *implicit* consolidation of correct responses. This theory is consistent with findings that populations benefiting most from errorless learning tend to display profiles of spared implicit memory alongside impaired explicit memory (e.g. Faulkner and Foster, 2002; Kuzis et al, 1999) There is also that evidence errorless learning promotes implicit knowledge in neurologically healthy participants (e.g. Maxwell, 2001), However, there is ongoing debate around the precise cognitive mechanisms sub-serving errorless learning, and there is some evidence for a role of spared explicit processing (e.g Kessels et al, 2005)

The majority of studies supporting the efficacy of errorless learning in ABI have been based around verbal learning. For example, Baddeley and Wilson (1994) studied word-stem completion, Hunkin et al (1998) studied list learning, Winter et al (1999) studied name learning, and Glisky and Delaney (1996) studied the learning of factual statements. Some studies have successfully applied errorless learning to vocational domains, for example Andrewes (1999) trained a participant in a filing task via errorless learning. Very few studies have assessed errorless route learning for people with acquired brain injury, although Brooks (1999) had great success in helping a severely amnesic patient to learn routes via a virtual simulation of those routes, in which he employed aspects of errorless learning. However, as the focus of the study was on virtual reality, and there was no controlled errorful comparison condition, it is not possible to unequivocally attribute the success to the errorless aspects of learning per se.

Evans et al (2000) failed to demonstrate an advantage for errorless techniques in two different route learning tasks, actually finding a benefit for trial-and-error learning in one. They note that lack of success of errorless learning could be due to ceiling effects, or due to a lack of participant involvement in errorless conditions. In addition, perhaps as a result of the difficulties surrounding running route learning studies described in the introduction, the tasks employed were quite abstracted from naturalistic 'route learning.' Participants learned routes from one piece of furniture to another around a two-dimensional sketch of a room, and learned routes across a grid of images. It is conceivable that a more realistic route-learning task may benefit more from errorless learning, particularly when one considers the fact that there is often a procedural element to real-world route learning route-learning (Allen and Willenborg, 1998, Garden, et al, 2002), and errorless learning has been linked with effective procedural knowledge acquisition (Maxwell, 2001).

The present study, therefore, sought to use an ecologically valid virtual environment to clarify whether errorless learning would be an effective technique (and more effective than errorful learning) for helping people with acquired brain injuries to learn routes.

4.2 Method

20 participants with acquired brain injuries were recruited from several sites (rehabilitation centres and hospitals) across the West Midlands, UK. Participants with language difficulties severe enough as to be likely to compromise understanding of task instructions were excluded, along with participants demonstrating visual neglect (assessed by the star cancellation task from the Behavioural Inattention Test, Wilson et al 1987). Participants had to have acquired their brain injury more than months prior to the experiment, and had to have a self-reported memory impairment (which was later confirmed via neuropsychometric testing).

4.2.1 Apparatus. The apparatus was the same as that used in Experiment 1.

4.2.2 Procedure. Participants learned one route in an errorless way, watching the entire route correctly through to completion three times before attempting to call out directions themselves. Another route, they learned in an errorful (or 'trial-and-error') way, being shown the correct route only once before being asked to take two practice attempts (during which errors typically occurred) at calling out the directions before the final, 'test' trial. When participants in the errorful practice trials called out an incorrect direction, they were taken a short distance down the incorrect street before being corrected by the experimenter. In this way, participants always saw each route through to completion three times before their final test, but in the errorless condition they saw it correctly each time, while in the errorful condition they saw the route plus, potentially, several incorrect detours along the way. The detours were corrected after a short time (5 seconds) in order to ensure that the two conditions were as similar as possible, apart from the errorful condition allowing participants to briefly leave the correct path. This was deemed important, to allow potential poor performance in the errorful condition to be attributed to the errors themselves, rather than simply to participants in this condition spending considerably longer walking around, experiencing more anxiety, or witnessing more junctions.

4.3 Results

A paired-samples t-test revealed that significantly fewer mistakes on the test route were made under errorless learning conditions than under errorful conditions ($t=2.631$, $df = 19$, $p = 0.016$). The mean number of errors made during the errorful test trial was 4.65 (S.D.2.35), compared with 3.4 (S.D. 1.54) during the errorless test trial.

4.4 Discussion / Conclusion

Participants with memory impairment resulting from acquired brain injury benefited more from errorless than from errorful methods when attempting to learn routes through a virtual town. This is consistent with findings from previous studies that have shown errorless learning to be an effective strategy in the learning of verbal information (e.g. Baddeley and Wilson, 1994). It is somewhat inconsistent, however, with findings from Evans et al's (2000) studies of errorless versus trial-and-error route learning. There are several possible reasons for this. It is possible that the virtual town task used in our study was more engaging than the paper-based route learning assessments used by Evans et al, because it was relatively novel and stimulus-rich. Boredom in errorless learning tasks, which are, by nature, very repetitive, can have an impact on performance, as participants may not pay full attention during learning trials. Tasks which are novel or interesting to complete may be more suited to errorless learning, therefore, as they are more able to hold participants' attention.

The fact that errorless learning was more effective than trial and error in the route learning task described here is theoretically important, as it is one of few studies to apply errorless methods beyond the realm of verbal learning. Further research into practical applications of errorless learning for everyday tasks could be potentially very fruitful.

The virtual reality town provided the ideal setting for studying errorless learning, as multiple and consistent trials can be presented without any distractions. There may be a place for virtual reality in training people with acquired brain injury in a range of diverse practical skills, using errorless principles.

5. EXPERIMENT 3: ERRORLESS LEARNING IN CONJUNCTION WITH NATURALISTIC STRATEGIES FOR ASSISTING PEOPLE WITH ACQUIRED BRAIN INJURIES IN ROUTE-LEARNING

A study in progress is exploring the relative effectiveness of two popular naturalistic strategies (for helping people with ABI to learn routes more effectively. The naturalistic strategies being combined with errorless learning are landmark use and cognitive map formation. A pilot study with neurologically healthy individuals, using the virtual town, found that the use of landmarks as memory cues correlated with the average accuracy of individuals' route-recall (rho = -.483, n = 40, p = 0.005). This is consistent with findings in the literature that efficient wayfinders tend to rely on landmarks to assist them (Kato and Takeuchi, 2003) The creation of a mental map of an environment is another key strategy that, although not as popular as landmark use, has also been seen to be employed by particularly good wayfinders (Prestopnik and Roskos-Eswoldsen, 2000, Billinghurst and Weghorst, 1995). Additional reasons for choosing these two strategies are that they are relatively easy to promote the use of, and that landmarks and maps could be argued to broadly tap verbal and visuo-spatial capacities, respectively.

It will be of theoretical interest to see whether errorless learning's efficacy can be improved by using it in conjunction with additional strategies, as the vast majority of errorless learning studies have not attempted to combine errorless learning with any other techniques. Discovering the efficacy of naturalistic strategies is also an important aim in itself. Anecdotal evidence from participants tested to date suggests that people favour tried-and-tested, familiar methods like the use of landmarks and maps, over novel 'rehabilitation' strategies like errorless learning.

It is predicted that naturalistic strategies will increase the benefit of errorless learning, through promoting deeper encoding of information. However, there is the possibility that the introduction of additional strategies will prove distracting, or will hinder the success of errorless learning by promoting more 'explicit' learning.

6. EXPERIMENT 4: PSYCHOMETRIC TEST PROFILES AND ROUTE LEARNING PERFORMANCE

An ongoing study, using participants from all of the patient studies described above, will assess the relationships between scores across a range of psychometric tests and route learning performance.

A battery of psychometric tests was selected based upon findings from previous route-learning studies. 6 tests, that have been previously found to correlate with measures of route learning performance, were chosen.

The tests that were chosen were as follows:

- The Block Design subtest of the Wechsler Adult Intelligence Scale - Revised (Wechsler, 1981). A test in which participants use 4-9 small blocks with two red, two white, and two split-coloured faces to reproduce patterns of increasing complexity. This test was selected because performance has been seen to correlate with route-learning performance on a real life driving task (Uc et al 2004a) and to predict accuracy on a landmark positioning task (Cutmore et al, 2000).
- The Money Road Map Test of Direction Sense (Money, 1965). In this test, the experimenter traces a route marked around a simple road map, and participant calls out the direction to take at each junction, without rotating the map. Performance on this egocentric mental rotation task has been shown to correlate with performance on a virtual reality maze task (Moffat et al, 1998).
- The Controlled Oral Word Association Task (Benton and Hamsher, 1976) involves the participant generating as many words beginning with a given letter as possible, within a minute. Moffat et al (1998) found that females' performance correlated with maze learning ability, and Uc et al (2004a) found performance of people with Alzheimer's dementia correlated with route learning performance on a driving task.
- The Rey Osterrieth Complex Figure Task (Rey, 1941, Osterrieth, 1944), in which participants copy, reproduce immediately, and reproduce after a 30-minute delay, a complex geometric design. Scores on this test correlated inversely with the number of wrong turns taken by participants (with Alzheimer's dementia and with acquired brain injury) in a real-life driving task (Uc et al, 2004a,b)
- The List-learning subtest of the adult memory and information processing battery [AMIPB] (Coughlan & Hollows, 1985).In this test participants are asked to learn a list of 15 words called out to them over 5 trials. The number of incorrect turns taken by participants with Alzheimer's dementia and

with acquired brain injury correlated inversely with scores on a very similar test, the Rey Auditory Verbal Learning Test (Rey, 1964).

- Digit Span and 'Spatial Span' subtests of the Wechsler Adult Intelligence Scale - III (Wechsler, 1981) were selected as tests of working memory. They involve, memorising and repeating increasingly long sequences of numbers and sequences of block-taps, respectively. Working memory is implicated in route learning, as increasing working memory load through allocating concurrent tasks appears to impair route learning. (Garden et al, 2002)

Statistical analyses will be carried out in order to assess:

a. Whether there is a correlation between route learning performance and any of the psychometric tests described.
b. Whether the advantage of errorless over errorful learning varies in line with participants' degree of impairment on certain psychometric tests.

If sufficient numbers of participants can be recruited, analyses will also be carried out to assess whether psychometric test performance predicts how effective landmark and map strategies are across participants. For example, do participants scoring high on Rey Complex Figure perform particularly well in the map strategy condition, because they have good visuo-spatial memory, or does verbal memory score correlate with performance in the landmark condition, with people tending to memorise landmarks verbally?

If relationships between psychometric tests scores and success of specific wayfinding strategies can be identified, by this or future studies, the practical applications are exciting. As part of rehabilitation programmes, administration of short neuropsychometric tests could potentially allow for patients with route learning difficulties to be provided with the most appropriate learning strategy for them.

7. GENERAL CONCLUSIONS

There has been great success in the use of virtual reality in route learning studies in general (e.g. Darken and Banker, 1998), but few studies of patients with acquired brain injury have used simulated environments to examine route learning. The present series of experiments demonstrates that virtual environments can be put to effective use in investigating rehabilitation strategies for people with acquired brain injuries. Furthermore, the ecological validity of the virtual reality town employed, demonstrated by comparing performance on a real and virtual version of a route learning task, shows that non-immersive, relatively inexpensive equipment can be a valid and productive research tool.

The present studies used the virtual reality town purely as a testing environment in which to compare various strategies. Many studies have striven to demonstrate the transfer of knowledge gained in a virtual environment to a real-world setting. While this is obviously a valid concern, the present study demonstrated the usefulness of the virtual world as a stand-alone arena for research. It is not essential that the virtual tasks completed in the present studies impact upon people's everyday route learning performance. The value of the virtual reality town was that it served as a controlled environment in which to address research questions. This use of virtual reality may be open to further exploitation in the future.

Having said this, a weakness of the studies is that it cannot be claimed, unequivocally, that the findings from the virtual town will apply to real-world route learning. There is a possibility that some feature specific to the virtual environment influenced performance in such a way that results would not apply in the real world. Real-world replications of the studies conducted in the virtual world would be the only sure way to reject this possibility, and future research addressing this question would be valuable. However, findings from the pilot study of good equivalence between performance and strategy use across the virtual town and the real world are reassuring.

Overall, virtual reality appears to be a very promising tool for researching route learning, and may be especially useful when assessing individuals with mobility difficulties or cognitive difficulties for whom route-learning in the real world may be impractical. Virtual reality is particularly well suited to the consistent repetition of learning material required when studying errorless learning, and it is suggested that more general applications of virtual reality for errorless learning research may be fruitful.

8. REFERENCES

G L Allen and J L Willenborg (1998), The need for controlled information processing in the visual acquisition of route knowledge, *Journal of Environmental Psychology*, **18,** 4, pp.419-427.

D Andrewes (1999), The Work Rehabilitation of a Herpes Simplex Encephalitis Patient with Anterograde Amnesia, *Neuropsychological Rehabilitation,* **9,** 1, pp77-99

A Baddeley and B A Wilson (1994), When implicit learning fails: amnesia and the problem of error elimination, *Neuropsychologia*, **32,** 1, pp53-68

A L Benton and K de S Hamsher (1989) *Multilingual aphasia examination*, Iowa City, Iowa: AJA Associates.

M Billinghurst and S Weghorst (1995), The use of sketch maps to measure cognitive maps of virtual environments, *Virtual Reality Annual International Symposium,* North Carolina, USA.

B M Brooks (1999), Route Learning in a Case of Amnesia: A Preliminary Investigation into the Efficacy of Training in a Virtual Environment, *Neuropsychological Rehabilitation,* **9,** 1, pp.63-76

A K Coughlan and A K Hollows (1985), In A. K. Coughlan (Ed.), *The adult memory and information processing battery*. Leeds: St James' University Hospital.

T R H Cutmore, T J Hine, K J Maberly, N M Langford and G Hawgood (2000), Cognitive and gender factors influencing navigation in a virtual environment, *International Journal of Human-Computer Studies,* **53**, pp223-259

R P Darken and W P Banker (1998), Navigating in Natural Environments: A Virtual Environment Training Transfer Study, *Virtual Reality Annual International Symposium*, Los Alamitos, CA, pp. 12-19.

R P Darken and B Peterson (2001) Spatial Orientation, Wayfinding, and Representation.In *Handbook of Virtual Environment Technology.* (Stanney, K. Ed.), Erlbaum, Mahwah, pp. 1-22.

J J Evans, B A Wilson, U Schuri, J Andrade, A Baddeley, O Bruna, T Canavan, S Della Sala, R Green, R Laaksonen, L Lorenzi and I Taussik (2000), A Comparison of 'Errorless' and 'Trial-and-error' Learning Methods for Teaching Individuals with Acquired Memory Deficits, *Neuropsychological Rehabilitation,* **10,** 1, pp67-101

D Faulkner and J K Foster (2002), The Decoupling of "Explicit" and "Implicit" Processing in Neuropsychological Disorders. Insights Into the Neural Basis of Consciousness? *Psyche,* **8,** Available from http://psyche.cs.monash.edu.au/v8/psyche-8-02-faulkner.html.

S Garden, C Cornoldi and R H Logie (2002), Visuo-spatial working memory in navigation, *Applied Cognitive Psychology,* **16,** 1, pp.35-50.

E L Glisky and S M Delaney (1996) Implicit memory and new semantic learning in posttraumatic amnesia, *Journal of Head Trauma Rehabilitation,* **11,** pp. 31-42.

G Grön, A P Wunderlich, M Spitzer, R Tomczak & M W Riepe (2000), Brain activation during human navigation: gender-different neural networks as substrate of performance, *Nature Neuroscience*, **3**, pp 404 – 408

S Haq, A Hill and A Pramanik (2005) Comparison of Configurational, Wayfinding and Cognitive Correlates in Real and Virtual Settings, *Proceedings of the 5th International Space Syntax Symposium,* Delft, pp.387-406

T Hartley, E A Maguire, H J Spiers and N Burgess (2003), The Well-Worn Route and the Path Less Traveled: Distinct Neural Bases of Route Following and Wayfinding in Humans, *Neuron,* **37,** pp.877-888.

N M Hunkin, E J Squires, A J Parkin, J A Tidy (1998) Are the benefits of errorless learning dependent on implicit memory? *Neuropsychologia*, **36,** 1, pp.25-36.

Y Kato and Y Takeuchi (2003), Individual differences in wayfinding strategies, *Journal of Environmental Psychology,* **23,** 2, pp. 171-188

R P C Kessels, S T Boekhorst and A Postma (2005), The contribution of implicit and explicit memory to the effects of errorless learning: A comparison between young and older adults, *Journal of the International Neuropsychological Society,* **11**, pp.144-151.

K C Kirasic, G L Allen and A W Siegel (1984), Expression of configurational knowledge of large-scale environments: Students performance of cognitive tasks, *Environment and Behavior,* **16**, *6*, pp. 687-712.

G Koh, T E van Wiegand, R L Garnett, N I Durlach, B Shinn-Cunningham (1999) Use of Virtual Environments for Acquiring Configurational Knowledge about Specific Real-World Spaces: I. Preliminary Experiment, *Presence,* **8,** 6, pp.632-656

G Kuzis, L Sabe, C Tiberti, M Merello, R Leiguarda and S E Starkstein (1999), Explicit and implicit learning in patients with Alzheimer disease and Parkinson disease with dementia, *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, **12**, pp.265-269.

J P Maxwell (2001), The implicit benefit of learning without errors, *The Quarterly Journal of Experimental Psychology: Section A,* **54,** 4, pp.1049-1068

S D Moffat, E Hampson and M Hatzipantelis, (1998), *Evolution and Human Behavior*, **19,** pp. 73–87

J Money (1976) A Standardized Road-map Test of Direction Sense, Academic Therapy Publications San Raphael, CA.

P A Osterrieth (1944), Le test de copie d'une figure complexe, *Archives de Psychologie,* **30, pp.** 206–356.

J L Prestopnik and B Roskos-Eswoldsen (2000) The relations amongst wayfinding strategy use, sense of direction, sex, familiarity and wayfinding ability, *Journal of Environmental Psychology,* **20,** 2, pp. 177-191.

A Rey (1941), Psychological examination of traumatic encephalopathy, *Archives de Psychologie,* **28,** pp. 286–340.

E Y Uc, M Rizzo, S W Anderson, Q Shi and J D Dawson (2004a) Driver route-following and safety errors in early Alzheimer disease, *Neurology,* 63, pp832-837

E Y Uc, M Rizzo, S W Anderson, Q Shi and J D Dawson (2004b) Driver route-following and safety errors in stroke, TRB 2004 Annual Meeting CD-ROM.

M Sidman and L T Stoddard (1967), The effectiveness of fading in programming a simultaneous form discrimination for retarded children, *Journal of the Experimental Analysis of Behavior*, **10**, pp.3-15.

R W Skelton, C M Bukach, H E Laurance, K G F Thomas, J W Jacobs (2000), Humans With Traumatic Brain Injuries Show Place-Learning Deficits in Computer-Generated Virtual Space, *Journal of Clinical and Experimental Neuropsychology,* **22,** 2, pp157-175

D Wechsler (1997) WAIS-III: *Administration and Scoring Manual: Wechsler Adult Intelligence Scale,* Psychological Corporation.

B A Wilson, J Cockburn and P W Halligan (1987), *Behavioural Inattention Test*, Harcourt Assessment.

J Winter and N M Hunkin (1999) Re-learning in Alzheimer's disease, *International Journal of Geriatric Psychiatry,* **14,** 11, pp. 988-990.

Assisting the mobilization through subway networks by users with visual disabilities

J H Sánchez, M A Sáenz

Department of Computer Science, University of Chile,
Blanco Encalada 2120, Santiago, CHILE

{jsanchez, msaenz}@ dcc.uchile.cl

www.dcc.uchile.cl/~jsanchez

ABSTRACT

We introduce AudioMetro, application software for blind users that represents a subway system in a desktop computer to assist mobilization and orientation in a subway network. A user can organize and prepare a travel by using the software before riding the subway. Conclusions of the usability study revealed the critical importance of using key interface elements, such as audio-based hierarchy menu, travel simulation, and information about the subway network, stations and their surroundings. Cognitive study results show an advance in the development of mobility skills needed for using the subway system which represent a contribution for a much more integral development of blind users and one-step towards social integration and inclusion.

1. INTRODUCTION

Biological beings have the natural need of pursuing autonomy and effective ways of mobilizing. Human beings depend on the particular qualities of each person in order to mobilize and orient in the surrounding environment. Users with visual disabilities have limited capacity to learn autonomously and effectively diverse ways of mobilizing. Actually the ways they use to mobilize are different to other people mainly because sighted persons understand the world through the vision. People with visual disabilities have to use orientation and mobility techniques to learn to recognize the environment, the spatial relationships between objects and themselves, and to recognize tactile, auditory, and olfactory stimuli. Once blind people acquire the ability to use orientation and mobility techniques they increase their capacity to mobilize effectively.

Through discrimination and localization of audio sources the sense of hearing convey information to people with visual disabilities becoming one of the most important senses to perceive the surrounding world. Along with this, the sense of hearing favors to know reference points, to perceive description of some places, to develop the abstraction and concentration capacity, and to create mental schemas of diverse environments (Sánchez and Sáenz, 2005a).

In order to navigate through real environments blind users take higher risks than sighted users because the use of the cane cannot help them identify all objects into the space. For this reason, it is necessary to provide cues to users with visual disabilities to get more reliable information. Providing too much information to users in an unnecessary way may have a counter effect and thus confusing users (Koruda et al, 2002; Lahav and Mioduser, 2004).

The use of sound to support graphical interfaces is a fundamental requirement in diverse tools for sighted users. Technology advancement allows us to obtain better results in sound representations of different environments, such as the greater and better immersion of users into virtual environments. For example, entertainment and communication requirements demand both the use of better quality sounds. Blind users need audio-based interfaces because they use the sense of hearing as the chief source of awareness and knowledge for learning purposes (Sánchez 2006a). Sound conveys information to identify points of reference, receive complete descriptions of places, and create mental models of the physical environment.

Blind navigation is based mainly on a perimeter exploration strategy. They walk along the closest outline in order to reach a certain targeting point. However, diverse attempts exist to achieve a more straight movement from one point to another, or between objects (Lahav and Mioduser, 2004). Independent and safe

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

navigation skills are obtained by a combination of motor, sensorial, and cognitive skills. Several studies have been made to promote the teaching and support the orientation and mobility of visually impaired people through virtual environments based on stereo or quadraphonic sound. Some of these studies present significant results on how visually impaired users are capable of mentally represent the navigated virtual environment as the first step in order to achieve a better movement (Lahav and Mioduser, 2004; Sánchez and Sáenz, 2005a; Westin 2004). Also, these studies show the importance of the utilization of haptic interfaces combined with sound to achieve the expected navigation in the virtual environment. (Lahav and Mioduser, 2004; Sánchez and Sáenz, 2005a; Westin 2004; Kapić 2003).

Visually impaired people need to describe their navigation routes with much more detail than sighted users. Specifically, blind people use objects like walls and railings to determine their direction, with special sounds and smells, such as water sources or bakeries. Relevant information can be obtained by the floor configuration and the changes present in it (Kapić 2003).

Interacting with different objects in real world distributed throughout the environment requires knowing and mentally representing the places walked or to be walked. Several studies have focused on the virtual representation of images, graphs, and textures, through haptic interfaces (Brewster 2002; Ramloll et al, 2001; Sjöström 2001). All these studies obtained important gains regarding the identification of the represented objects. However, most of them aimed at the need of always requiring an acoustic support of the object's representation, either as an introduction or to make clear certain aspects of it. The fact that the haptic sense is slower and less efficient than the vision channel has to be considered. Actually, that is why it is so important to support and complement it by means of the hearing sense.

Audio-based interfaces may have a speech or a non-speech audio. Non-speech audio allows the improvement of interaction and the presentation of information in different devices. This helps blind users to be concentrated on navigating the physical environment by being aware that information is perceived by their ears (Sánchez 2006a). Users with visual disabilities evidence many mobility problems. One of them is the particular case of mobilizing through the subway network. In Santiago city, the subway is a public transportation on way of expansion with better comfort and minor travel duration time than other type of public transportation, then it is very important that blind users can use these benefits more fully.

2. RESEARCH PROBLEM

The lack of knowledge about some aspects of the subway network is the principal mobilization and orientation difficulty for subway users. To obtain a whole knowledge of the subway network a person has to pass through three levels of knowledge: (A) Conceptual, to know subway network concepts such as subway lines, stations, and platforms. This level refers mainly to basic and generic concepts in a subway network in any city of the world. Once the concepts are learned, the user continues to the next level; (B) Knowledge, to know specific information of a particular subway network, such as the subway station name, surroundings, location, and lines names that identify a route or path and what type of station is it (transfer and local), and (C) Articulation, to utilize different concepts and knowledge learned by the user for an efficient use of the subway network. It includes mainly the planning, estimation, cost, and spatial orientation of the travel. When the three levels are accomplished, the user must also be able to master the orientation into the space to mobilize independently. When users lose spatial orientation, their autonomy is decreased.

Users with or without impairments that currently utilize the subway network present different spatial and knowledge problems. Spatial problems are easier to solve for sighted users because they can visually order and classify the physical environment by taking advantage of the visual memory thanks to the use of visual references such as stores, colours, and signals of the subway network. Users with visual disabilities orient themselves through sound interpreting the surrounding world and learning to localize sounds that serve as signals for orienting and moving with greater autonomy.

In the real world sounds are not necessarily fixed, nevertheless visual references are fixed. For this reason, it is necessary to have better tools that support to solve orientation problems. Blind people can face knowledge problems because the vision cannot be used as an imitation channel to acquire knowledge. The affordance of elements is a very important property detected by the vision.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

3. PURPOSE OF THE STUDY

The main purpose of this study was to create, improve, and evaluate audio-based software to stimulate and/or develop tempo-spatial sensory and cognitive skills for mobilization and orientation in people with visual disabilities. These skills allow users to have safe and independent movements facilitating their autonomy when using different transportation media such as the subway.

In this study we also consider the implementation of a usability study during the design and development of the software. In order to know the cognitive impact of the use of the software, learners solved cognitive tasks specially designed and created for this purpose.

4. MODEL

AudioMetro is based on a model that incorporates diverse functionalities in order to evaluate and identify a correct feedback, and establish differences in the development of educational software for blind or sighted people (Sánchez and Baloian, 2005b). AudioMetro contains a metaphor that represents a simulation of a subway travel through a wagon. Travels are developed in a logical way because the software does not consider spatial representations of virtual spaces. The metaphor considers notions of consecutive stations, transfer stations, and terminal stations.

In the particular case of the subway network in Santiago, Chile, the subway wagon travels between two stations from one extreme to the other through a specific line that covers both directions. The stations have two platforms, one on each side of the rail. In these platforms, passengers wait for subway wagons at specific directions. Transfer stations consist of different levels where each level has a specific line and each line crosses the other, to allow associating each line to a level.

In each AudioMetro session before every simulation, the user must previously choose the initial and final station of the travel. In order to know at any moment the game state and to allow all necessary functionalities to execute in a virtual travel, AudioMetro utilizes techniques of object-oriented programming to model the stations, lines, network lines, and travels. This model calculates the optimal route from the current station to the final station in order to both, user and software, take strategic decisions during simulations. AudioMetro shows information to the user through an interactive menu hierarchy to transfer information in the three levels of knowledge. This menu hierarchy is presented through audio and can be complemented with the contextual help for interacting with the software. User actions are saved in a session log file for later analysis in order to evaluate the learning of the user by using the software.

The logic model of AudioMetro provides the possibility of representing any subway system, identifying the composing stations and lines by specifying boarding and terminal stations.

5. SOFTWARE DESCRIPTION

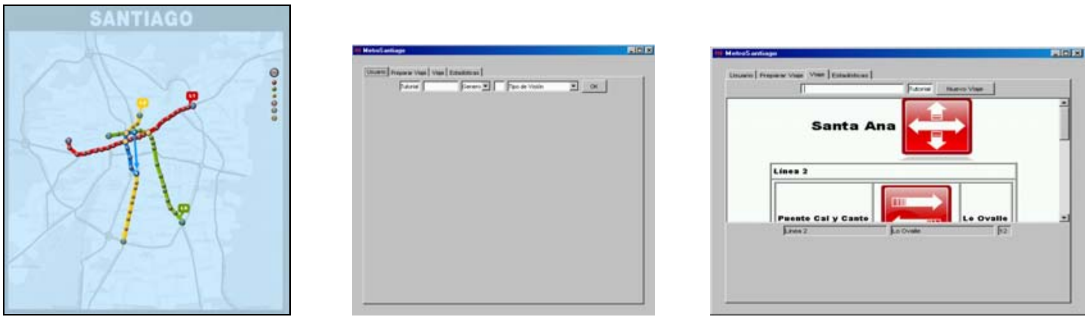
A work session with the software has two stages. The first stage consists in preparing a travel with a user defining starting and ending stations. In the second stage, the user virtually moves through the subway network, starting at the initial station.

At the beginning, the software does a random choosing of values for each of three main variables in a travel through the subway network. The three variables are: Subway line, travel platform (direction), and starting station (transfer or local).

Figure 1(A) shows a travel from the *Irarrázaval* station (line 5) to *Franklin* station (line 2) of the Santiago's subway network. The symbols used in figure 1 for the subway network are: Line 1 stations with red color, Line 2 stations with orange color, Line 5 stations with green color, transfer stations with yellow color, the way from the starting station to the ending station with light blue color, and the light blue arrow represents the direct way from the starting to the ending station.

AudioMetro was developed with Java using the Swing library. The software code is divided into four main packages: (A) *Metro*, defines the objects that represent a subway network; (B) *DomArea*, groups all classes related to XML documents management; (C) *Navegacion*, represents the navigation through menu hierarchies, feedback, sounds coordination, and others functionalities of the software, and (D) *Sonido*, groups all methods related to the use of sound.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3



(a) *(b)* *(c)*

Figure 1. *(a) Travel from Irarrázaval station to Franklin station. (b) AudioMetro user interface (c) AudioMetro´s trip interface.*

5.1 User-Software Interaction

The graphical objects of AudioMetro interface are organized into three panels. When starting the software the user must input personal information in the first panel. This information is saved into the session registry for a later analysis. Then, in the second panel, the user chooses both starting and ending stations. At last, in the third panel, the user explores the subway network and associated contents to the three levels of knowledge. The main graphical objects are text boxes, combo boxes and buttons, using the keyboard as input. The software provides the necessary audio feedback for users with visual disabilities in order to use these visual components. In the case of navigation using menu hierarchies, at the third panel, the user interacts exclusively with a text box that captures keyboard events without changing the focus between different graphic components (see Figure 1(B),(C)).

6. USABILITY EVALUATION

6.1 Participants

The sample consisted of 7 Chilean users with visual disabilities, 5 males and 2 females, ages 15 to 32, which attend the rehabilitation level of "Santa Lucía" School for the Blind in Santiago, Chile. Three of them have low vision and four of them with total blindness. All of them have acquired blindness. Two special education teachers and one usability expert also participated in the study.

6.2 Methodology

Usability testing with end-users had five stages: 1. **Introduction**. The user received explanations of the purpose of testing and how to use the keyboard to interact with the application. Teachers mediated to help users to orient when using the keyboard. 2. **Software interaction**. Users had to use the software. Each participant performed between two and three travels and according to their needs, teachers could assist them for better orientation. At the same time we wrote anecdotic records with key data and observations of the user's interaction with the software (see Figure 2). 3. **Application of usability questionnaires**. The user answered questions concerned with software interaction asked by special education teachers. 4. **Reports of sessions**. Each session was photographed to register the child's behavior during interaction. All data was registered to get comments and suggestions to improve the software navigation. 5. **Design and redesign**. According to the comments and observations, the software was redesigned and some new functions were designed.

Figure 2. *Users interacting with AudioMetro.*

6.3 Instruments

We used three evaluation instruments**: 1. End-user questionnarie** was applied at the end of the usability sessions. It is basically a software acceptance test and consists of 18 closed statements with an answer scale of 1 to 10 points. It also contains 5 open questions. This questionnaire was read by teacher and answered by the user. 2. **Anecdotic record** was used to record information obtained through observation when the user interacted with software. 3. **Automatic record** was data automatically stored in XML format and registered the keystrokes and stations used, and the duration time of each action.

6.4 Procedure

The usability testing was implemented in 5 sessions between March and June of the year 2005, following three stages. In the first stage, users tested early prototypes of the software during interaction. The objective was to have an initial feedback about the sounds of the software in order to have early in the implementation phase, the information necessary to guide the redesign of the interface. The second stage was implemented after we processed the data from the first testing and redesigned and improved the prototype. Therefore at this stage we had a more advanced prototype. The results of the second testing allowed to guide the final design of the interface. Finally, the third stage was applied to the final results of software usability. The aspects evaluated were: motivation, importance, relevance of use, and the sound of the software.

6.5 Results

Figure 3 shows the results of the final usability test. The statements used to evaluate the motivation of the software were: “I like AudioMetro”, “The software is pleasant”, “The software is challenging”, “AudioMetro makes me to be active”, “I would like to play again the software” and, “AudioMetro is motivating”. For evaluating the importance of the software the statements were: “I recommend you this software to other people”, “I learned with AudioMetro”, and “The software allowed me to know new things”. For the relevance of use we considered the following statements: “I felt controlling AudioMetro”, “The software is easy to use”, and “AudioMetro adapts to my pace”. For sound evaluation, the sentences were: “I like the sounds of the software”, “The sounds are clearly identified”, and “The sounds of AudioMetro convey information”.

Users with residual vision showed a greater motivation and acceptance of the software from the beginning of testing to the end. Blind users did not show the same motivation, their scores were not significantly different in comparison to users with residual vision, reaching in the last test scores of 8.8 out of 10 points. One of the most relevant motivations for users was the assigned value to the tool to support the autonomous planning of routes and travels. These actions are currently made with the help of other person, especially in the beginning of independent displacements. All users assigned high importance to the help of tools such as AudioMetro. This is reflected on the statement “The sounds of AudioMetro convey information”, which always got the highest acceptance score. The easiness of use of the software is practically the same for both type of users, reflecting a good design of the tool. To users with residual vision, sound was improved during the implementation phase reaching the highest acceptance rate in the final testing. For blind users the final evaluation reached a score of 8.8 out of 10 points, showing a high acceptance of the sounds used in the software. The highest score was obtained in the statement “The sounds of AudioMetro convey information” (see Figure 3).

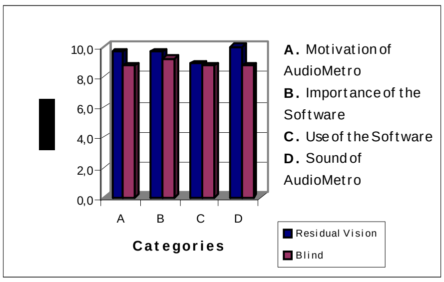


Figure 3. *Usability testing results.*

7. COGNITIVE EVALUATION

6.1 Participants

The study was designed with a sample of 10 users from the rehabilitation level of the School for the Blind “Santa Lucía” from Santiago, Chile, ages ranging from 20 to 32 years old. This group was divided into two equal subgroups: A control subgroup, which did not have any interaction with the software, and an experimental group. The subgroups were created so that the age and the number of people with residual vision were balanced.

6.2 Concrete Materials

To observe the mental image users created of the geographic distribution of the Santiago's subway we worked with two types of materials: 1. Cardboard base and LEGO© bricks. This allowed the sample to represent details of stations of the Santiago Subway network by building them with the different bricks, which varied according to whether it was a local or a transfer station; and 2. Cardboard base with stakes. It allowed users to represent the distribution of the Santiago subway network lines, by joining different stakes through rubber bands.

6.3 Methodology

The activities were divided into four steps: 1. **Pre-Test Application.** Users had to make a real travel route in the Santiago subway to evaluate and register the users' actions and their control in an independent journey. *2.* **Software or Class Interaction.** The experimental group knew and interacted with AudioMetro software by means of significant activities to assimilate, generalize, and take to reality their learning. The control group obtained similar information that is embedded in the software but through a regular lecture session in a classroom setting. 3. **Cognitive Tasks.** There were four cognitive tasks solved by users to observe the development of the following skills: to know and apply tempo-spatial concepts, to make an efficient use of senso-perceptive organs, and to select and apply concepts provided by the software for independent movements through the Santiago subway network. 4. **Post-test Application.** In the same way as in the pre-test, users made a real travel route in the Santiago subway to evaluate and register the learners' actions and control in an independent journey.

6.4 Instruments

There were three instruments for cognitive evaluation: 1. **Pre and Post-test**, Specific Route Displacement test (SRD), a particular route was elaborated for testing the participants. We evaluated their performance using an appreciation scale. In order to elaborate both the route and the evaluation range, we considered the use of certain orientation and mobility skills, the same route for everyone, departure hour, number of transfers, number of stations, visual degree, and knowledge with the real displacement through the Santiago subway. 2. **Cognitive task evaluation guidelines**, estimation evaluation guidelines were created for each cognitive task to observe and register orientation and mobility skills to be developed, stimulated or enhanced by solving the tasks and participating in the activities. 3. **Registry Graphs**, the software registered and graphed actions taken by learners during their interaction.

6.5 Procedure

The cognitive testing was carried out during four months. As the stages of the applied methodology were fulfilled, and depending on the subgroup to which each user belonged, the cognitive testing was implemented in different places. For both subgroups, pre-test and the post-test were applied on site. Experimental group users interacted with the software in the computer room whereas the control group users attended classes in a conventional classroom. All users in a classroom performed tasks with concrete materials. In order to avoid distortions in the results of the cognitive testing, only the experimental group interacted with AudioMetro and the entire group used the same version of the software.

6.6 Results

Results are divided in three areas or domains: Behaviors, skills, and competences. The first one refers to the specific handling of techniques with the white cane, which correspond to the orientation and mobility program applied to environments of medium complexity. The skill area includes aspects needed to make independent movements. Finally, the competence domain shows the development level that each user performed in each one of the described skills. Figure 4 displays comparative results of pre-test and post-test.

The results show that the greater improvement took place in the competence domain, with an increase of a 20% out of the total, followed by the skills domain, with a 16%, and the behavior domain with just a 6%. Nevertheless all areas obtained improvements after using the software and cognitive tasks. The variation in the behavior domain of the sample is the one that obtained the slighter difference in the scores. Learners required knowing certain concepts related to this transportation media and knowing how to use them, such as: climbing and descending stairs to reach the stations, navigating in their interiors identifying ticket offices, ticket gates, platforms, texture recognition, and remaining in a safe place when getting on and off the wagons. These behaviors did not show great enhancements since they are prerequisite of a relatively independent movement. Nevertheless, improvements in learning the stations, sequential order and its relation

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

to the corresponding line, could be observed. The actions corresponding to behaviors require certain skills such as spatial orientation, observation of the environment, and acoustic information classification. These aspects were remarkably enhanced after using the software and applying the cognitive tasks, since the subjects could make a mental model of the space distribution closer to reality and were able to move with more security and autonomy.

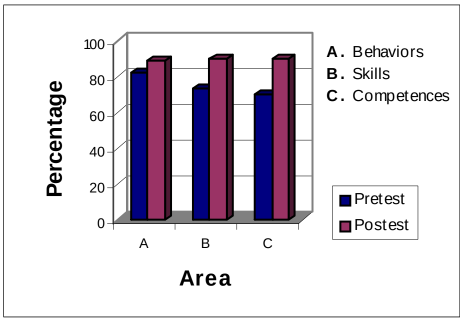


Figure 4. *Cognitive test results, pre-test and post-test comparison.*

During cognitive tasks each person had the opportunity to represent with concrete material the mental image of the subway, which helped the mediation of facilitators. This leaded us to modify the initial designs based on the experience, usefulness, and landmarks, to establish new concepts of orientation (north, south, east, west, northeast, northwest, southeast, etc.), that in spite of knowing them, they did not know their corresponding location in the space when applying these concepts.

User training by using AudioMetro software and supported by cognitive tasks, showed to be an important tool to encourage independent, safe, and significantly more integral navigation. This is verified in the comparative analysis of the results obtained by different users in the pre-test and post-test evaluations. This is also highlighted in the contrast between the time required by different users to follow a route assigned in the pre-test phase, and the time needed to take the same route after using the software, in the post-test phase. In all cases the time decreased in considerably way, almost in 20% average.

From the AudioMetro research, some questions emerged about the impact that a similar mobile-device oriented application would have on mobilization. For this reason, we later developed mBN (Mobile Blind Navigation), a pocketPC version of AudioMetro (Sánchez and Sáenz, 2006b). We believed that this mobile software does not have to be just a portable AudioMetro version. Rather, we had to redefine and study both the interface and the information displayed. We have to define the functionality to be exploited in the mobile version, as well as the cognitive impact. Although this new software aims to the same users, it is subtly different since it is oriented to a different context, but complements the scope of AudioMetro. The essential of the mobile version is that it can help to solve unexpected problems when mobilizing through the subway, which could hardly be anticipated in the PC version.

8. DISCUSSION

At a conceptual level, AudioMetro glossary allowed users to familiarize themselves with basic concepts of subway as a transportation source. The most important concepts (transfer stations, platforms, end stations, consecutive stations, and lines) were reinforced by the affordance created with travel simulations. When blind users interacted with the software, they were able to understand the subway network system as a means of transportation in the three levels of knowledge aforementioned without the need of using the subway with a guide companion. These trip simulations also supported the learning of specific knowledge of the subway network system of Santiago, by means of exploring the subway network names and characteristics of the stations, lines, surroundings, etc. Since these are simulations, they implicitly motivate the construction of virtual trips through the subway network.

AudioMetro was more useful for blind users in a pre-cane stage. We must remember that these users still did not possess the necessary autonomy to move throughout the subway network by themselves. Teachers valued AudioMetro as a support tool to their classes as it allowed them to apply, reinforce, and complement their mobility and orientation lessons. The usability test showed the importance of embedding in the software a glossary of concepts and surrounding information about stations. The anecdotic record permitted us to identify critical areas of the software, to debug feedback messages (voice and sounds), and to improve the software contents.

The variation in the mastering of behaviors by the user was generally the one that obtained the least indicators, which is explained by the fact that learners necessarily needed to know previously certain concepts related to this type of transportation and know how to use them. Among these concepts, there are activities such as going up and down stairs in order to have access to the stations, the ability to move inside the stations and identify ticket offices, turnstiles, and platforms, recognition of textures, and staying in a safe place when getting in and out the subway wagon. These behaviors did not show critical changes because they are pre-requisites for relatively independent movement. However, there was an improvement in the learning of the names of stations, their sequence, and their relation to a correspondent line. The execution of actions corresponding to the behaviors needs to master certain abilities such as spatial orientation, observation of the surroundings, and classification of auditory information. These skills were notable increased after using the software and performing cognitive tasks, because users were able to build a mental idea of the spatial distribution closer to reality and were able to move in a secure and autonomous way.

Finally, the most relevant results were observed in the domain of competencies of users that is explained by the information given by the software and its immediate transfer to reality that favored the rise of sensorial-perception information processing, problem solving without the need of a sighted guide, and a much more independent movement in and out of the subway network. This, without a doubt, represents a big contribution for a more integral development of blind users and one-step towards real social integration.

Acknowledgements: This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158. The authors would also like to thank Eduardo Maureira for programming the prototype version of AudioMetro.

9. REFERENCES

S A Brewster (2002) Visualization tools for blind people using multiple modalities. Disability and Rehabilitation Technology. 24 (11-12), pp 613-621.

T Kapić (2003) Indoor Navigation for Visually Impaired. A project realized in collaboration with NCCR-MICS. Julio – Octubre 2003.

T Koruda, H Sasaki, T Tateishi, K Maeda, Y Yasumuro, Y Manabe and K Chihara (2002), Walking aids based on wearable/ubiquitous computing – aiming at pedestrian's intelligent transport systems. Proceedings of the IV International Conference Disability, Virtual Reality & Associated Technologies, 2002 ICDVRAT 2002. Veszprém, Hungary, pp 117-122.

O Lahav and D Mioduser (2004), Blind Persons' Acquisition of Spatial Cognitive Mapping and Orientation Skills Supported by Virtual Environment. In Proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2004, Oxford, UK, 2004. pp. 131-138.

R Ramloll, W Yu, B Riedel and S Brewster (2001) Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users. In Proceedings of BCS IHM-HCI 2001 (Lille, France), Springer, pp 515-530.

J Sánchez and M Sáenz (2005a), Sound Immersed Virtual Environments for Blind Children. Proceedings of The Fourth International Workshop on Virtual Rehabilitation IWVR '05. Catalina Island, California, USA. September 19-21, 2005, pp. 192-202

J Sánchez and N Baloian (2005b). Modeling Audio-Based Virtual Environments for Children with Visual Disabilities. Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications ED-MEDIA 2005. Montreal, Canadá. June 27 - July 2, 2005.

J Sánchez (2006a). Blind Children Centered Technology. Human Centered Technology Workshop 2006. Pori, Finland , June 11-13, 2006, pp. 104-112

J Sánchez and M Sáenz (2006b). Subway Mobility Assistance Tools for Blind Users. 9 th ERCIM UI4ALL Workshop. LNCS (in press).

C Sjöström (2001) Designing Haptic Computer Interfaces for Blind People. Presented at ISSPA 2001, Kuala Lumpur, Malaysia, August 13 - 16, 2001.

T Westin (2004). Game accessibility case study: Terraformers – a real-time 3D graphic game. In Proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2004, Oxford, UK.

ICDVRAT 2006

Session VI. Medical Treatment and Home Based Rehabilitation

Chair: Pat Langdon

Using virtual reality for medical diagnosis, training and education

A H Al-khalifah[1], R J McCrindle[1], P M Sharkey[1] and V N Alexandrov[2]

[1]School of Systems Engineering, the University of Reading,
Whiteknights, Reading, RG6 6AY, United Kingdom

[2]Advanced Computing and Emerging Technologies Centre, the University of Reading,
Philip Lyle Building, Whiteknights, Reading, RG6 6BX, United Kingdom

a.h.al-khalifah@reading.ac.uk, r.j.mccrindle@ reading.ac.uk,
p.m.sharkey@ reading.ac.uk, v.n.alexandrov@ rdg.ac.uk

[1]*www.sse.reading.ac.uk,* [2]*acet.rdg.ac.uk*

ABSTRACT

In this paper we present a number of the immersive VR applications that we have developed during the past 18 months as a means of practically demonstrating the modelling approaches previously reported. The paper discusses the usefulness of the different approaches in assisting medical practitioners to diagnose and track conditions which might lead to impairment or disability, and how they can be used to train medical students to recognise such conditions or to undertake associated medical procedures. Initial findings of a survey of undertaken with medical practitioners as to the effectiveness of VR and in particular immersive models as diagnostic and training aids are also presented.

1. INTRODUCTION

To assist medical students in their understanding of anatomical structures or medical procedures, three-dimensional (3-D) virtual models can be used which offer a number of advantages over traditional 2-D images, plastic models and cadavers. For example, 2-D images have the inherent problems of limiting the view to one angle and obscuring important details; plastic models have the limitation of not presenting the fine detail; and cadavers are expensive to obtain and maintain, while their use may raise ethical and moral questions for some people. Digital 3-D models however can be accessed electronically and interacted with in order to enhance understanding. Immersive virtual reality devices such as CAVE-like environments (Cruz-Neira et al 1993) can further enhance this interaction and understanding as well as allowing collaboration to occur both locally and at a distance. General models for display within immersive environments can be generated using library data such that available from The National Library of Medicine's Visible Human Project (NLM 2006). Alternatively, data for individual patients can be displayed, obtained for example, from imaging scans.

At ICDVRAT 2004 (Al-khalifah and Roberts 2004) a number of modelling approaches for medical simulation were presented. In this paper we present a number of the immersive-based VR applications that we developed during the past 18 months in order to practically demonstrate such approaches. We consider the usefulness of these models to be in areas such as assisting medical personnel to diagnose debilitating conditions which might lead to impairment or disability, for monitoring improvement of a patient after therapy; and in training medical students to recognise particular conditions and other medical structures. With regard to this, initial findings of a survey of medical practitioners as to the effectiveness of VR and CAVE-like environments as medical training and diagnostic aids are presented.

2. VR MODELLING APPROACHES

A number of benefits can be gained through using models within immersive CAVE-like environments. For example, trainees, educators and/or practitioners do not need to be in the same physical environment to view and interact with the model. A demonstrator for example, can be in a remote location working via a virtual interface whilst sharing their model(s) with students or other practitioners who are located elsewhere in a CAVE-like display. Interaction with models in the immersive environment may also be more intuitive than

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

that with traditional 2-D models and they may also allow the user greater freedom of movement and ability to explore structures from multiple viewpoints. Practitioners can for example interact with the model(s) to explain anatomical structures or the symptoms of progressive illnesses such as glaucoma to the students. Time progressive models of an individual patient's brain scans could also be displayed to help practitioners monitor the effects of treatment and recovery of that patient over time, for example after suffering brain trauma from an accident.

Approaches such as polygonal, iso-surface and volumetric modelling techniques can be used to represent different types of medical conditions, procedures or structures. Decisions as to which approach is the most appropriate for a particular medical situation can be made by considering the properties of the different models.

Polygonal (surface) modelling focuses on creating abstract representations of structures and as such requires far less computational resources than other forms of rendering. Polygonal representations are artistic impressions of models created by the developer, composed of a set of polygons, and with accuracy and smoothness of the models governed by the number of polygons. This type of modelling allows developers to create structures according to their requirements in terms of shape and accuracy making them suitable for modelling human body structures and anatomical representations. Polygonal representations are fast to render, but this gain in speed comes at the cost of a loss in accuracy.

Iso-surface modelling (indirect volume rendering) is another common modeling approach in medical visualization and simulation implementations (Lorensen et al 1996; Gingins et al 1996; Kalra 1995). This procedure involves the generation of surfaces from volumetric data. The process follows a segmentation stage, which extracts the desired material surface from original volumetric data. Rendering speeds of such implementations are normally increased due to the dramatic reduction in data size. Surface texture mapping is another advantage that allows the addition of realistic images obtained from real organs to the surface of the model to enhance appearance. However, such high rendering speeds and aesthetic appearance of the models come at the price of a loss of internal data representations.

Volumetric modelling (volume rendering) is a visualisation technique (Kaufman et al 2000, 1993a, b; Gibson 1998), where 3-D models of human organs are generated from scanned images, such as Magnetic Resonance Imaging (MRI), Computer Tomography (CT), or Positron Emission Tomography (PET) scans. These images are then displayed as high quality voxel-based 3-D volumetric models. The main advantage of such a procedure is the ability to preserve original anatomical structures and details from original data. Its main drawback however is the need for powerful computational systems to handle the large quantities of data involved. Such conditions have serious implications on rendering speeds of the models and the required massive storage capacities to hold such data.

A number of medical applications have been developed to explore the issues associated with each modelling approach.

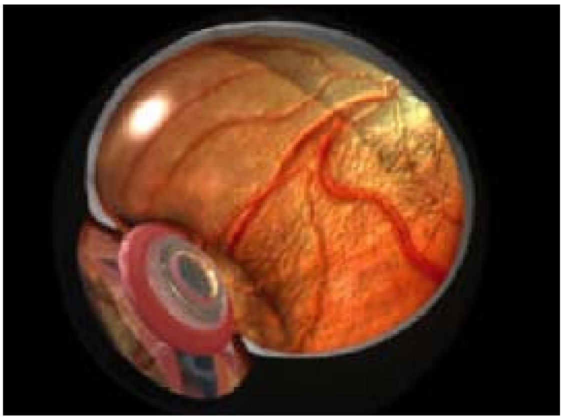
3. THE APPLICATIONS

Medical diagnosis, training and education activities can involve all of the above mentioned techniques, but they vary in the degree to which they can be effectively used and how they can be applied. Simulation, modelling and visualization are the main techniques that can be employed and a number of immersive virtual training applications have been developed, each one offering unique characteristics. In particular the 3-D models developed allow local and remote users to interact in real time with the models and for discussions to be shared amongst users. For demonstrative and educational purposes both polygonal and iso-surface modelling approaches were also adopted in this work. A number of the applications developed are briefly described below.

Polygonal modelling was used to develop the virtual eye shown in Figures 1 & 2. (Webb et al 2003). This application enables practitioners and educators to study the anatomy of the eye for educational purposes and also to observe how progressive illnesses such as glaucoma may develop over a period of time.

Polygonal modelling was also used to develop the hip replacement application (Figure 3) which can be used to demonstrate the anatomical structure of the hip and to allow students to practice the procedures that would be involved in carrying out such an operation.

Iso-surface modelling was used to create knee and brain applications for both anatomical demonstration purposes and diagnostic tools (see Figures 4 & 5). The particular advantage of both of these applications is that they represent accurate and realistic structures generated from real human body data scans. This allows practitioners to examine real organs of the human body and inspect possible traumas or deformities.

Figures 1 & 2. *The model of the eye. User interacting with the model (left). The model is clipped for better inspection (right). These images appear courtesy of [Webb, 2003].*

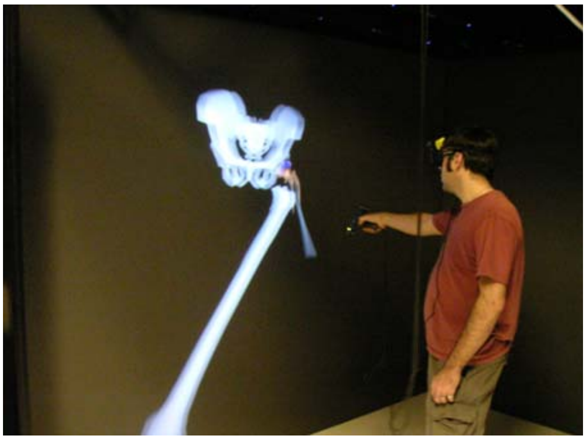
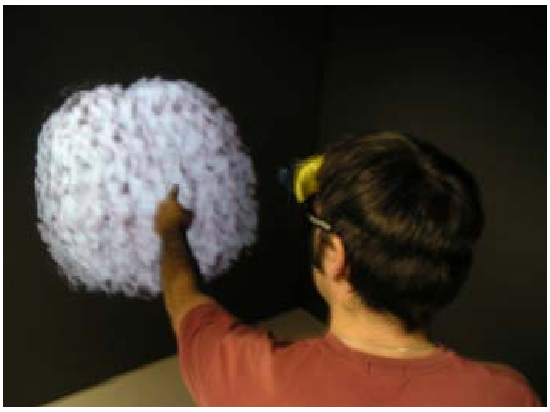


Figure 3. *User demonstrating a hip replacement procedure.*

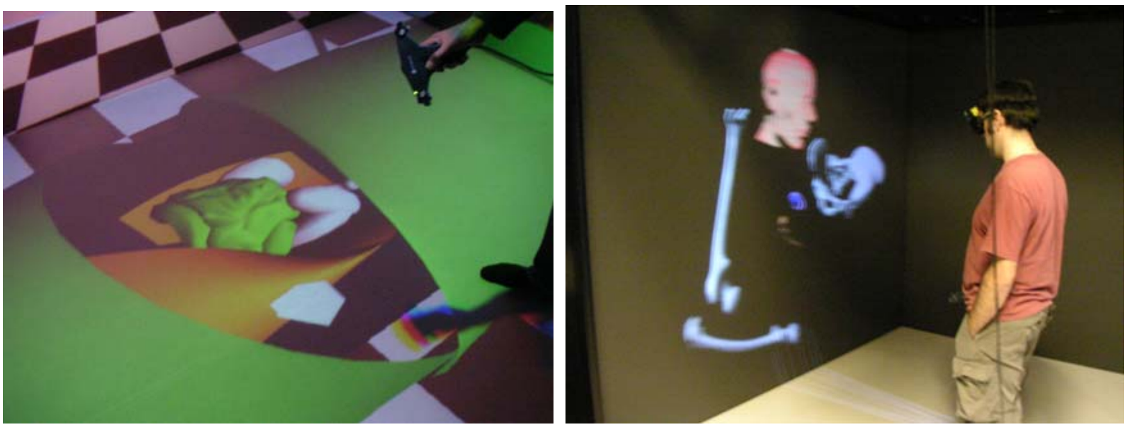
Figures 4 & 5. *An isosurface model of the knee (left) and isosurface model of the brain. Users inspecting and interacting with models.*

Human head and brain applications were also developed using volumetric modelling as diagnostic tools (see Figure 6). The National Library of Medicine's Visible Human Project (VHP) datasets were also displayed in some of our immersive applications including those of the male thorax as shown in Figure 7. It was felt important to ensure that this dataset could be modelled within the immersive environment as it forms a scientific archive of medical datasets for educators and professionals working in the field of medical visualization.

Figures 6 & 7. *User examining a volumetric model of the human head (left).User interacting with the thorax volumetric dataset from the Visible Human Project (right).*

For simulation purposes, open surgery applications was also considered in the development phase of this work. Demonstrations of basic deformation and open surgery procedures were implemented as shown in Figure 8. This includes the opening of the abdominal area and the inspection and interaction with organs such as the stomach and kidneys.

In order to compare different anatomical structures and study the human anatomy from different perspectives and levels of detail, a combined modelling approach was also adopted in order to create an educational application that demonstrates an anatomical structure of the human body using the three different modelling approaches (polygonal, iso-surface and volumetric) as shown in Figure 9.



Figures 8 & 9. *Internal parts of the human body exposed following the deformation of the abdominal cavity (left). User interacting with different human body parts modeled using different modelling techniques (right).*

Because collaborative education is a critical tool in distance and local learning, networked applications were also developed to allow collaborative tasks to occur such as the sharing of models and datasets. In this instance of an educator being in a remote location, working on a virtual interface whilst sharing this model and its constituent parts with students, the educator performing the actions on the model/models can be seen by the students as an avatar representation.

4. THE STUDY

In order to gain feedback on our applications and to ascertain the views of practitioners on the applicability of virtual reality for medical training, diagnosis, and surgery planning, two studies were conducted based on questionnaires and CAVE-based demonstrations.

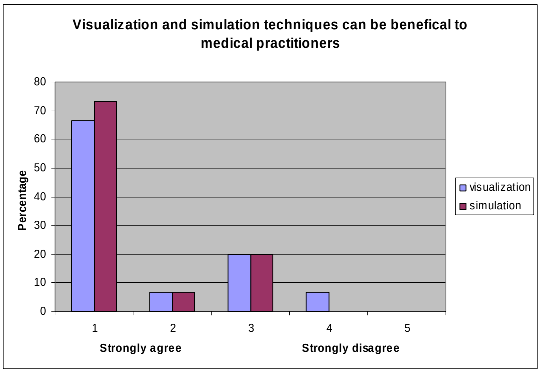
In the first study views from 16 medical professionals (Group A) including surgeons, general practitioners, medical consultants, clinical engineers, clinical skill directors and medical educators were obtained. The participants in this group were familiar with VR technology and were also given a 20 minute demonstration session inside the immersive display where they had the opportunity to interact with a number of our medical applications ranging from diagnostic (volumetric) to educational (polygonal and isosurface) models. Following this display they were given a questionnaire to complete.

In the second study feedback was sought from 12 medical practitioners (Group B) including medical consultants, general practitioners, researchers, surgeons, and medical educators who differed from Group A in that they had no prior knowledge or experience of virtual reality. Nor did this group receive an immersive demonstration of the applications although the questionnaire did have a number of related images attached.

The questionnaires covered issues relating to the applicability of VR technology in general, and the immersive display in particular, to medical education, visualization and simulation. Questions relating to the effectiveness of the immersive technology within the simulation and education processes compared to desktop devices were also considered but are not reported here.

5. RESULTS AND DISCUSSION

The partial outcome of these studies, comparing the perceptions of the two groups on the usefulness of VR to medical practitioners is outlined below:



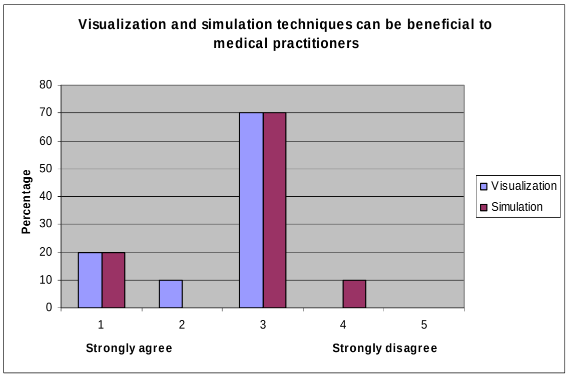


Figure 10 (Group A) & Figure 11 (Group B) – *are visualization and simulation techniques beneficial to medicine?*

Figures 10 and 11 give the overall responses from participants being asked if they considered whether visualization and simulation techniques could be of benefit to the medical profession. Here we note that Group A, who had experienced first hand the immersive VR application demonstrations responded much more favourably than those in Group B who had not seen the demonstrations.

Figures 12 and 13 show the breakdown of the usefulness of the different modelling techniques -volume, iso-surface and polygonal (surface) - for medical applications as perceived by the two groups. Once again participants in Group A tended to perceive the benefits of the different modelling approaches more strongly than those in Group B but even here very few respondents thought that they would be of no use at all. Both groups tended towards volumetric modelling as being the approach that offered the greatest potential for use in medical applications possibly due to the amount of data that can be displayed using this technique.

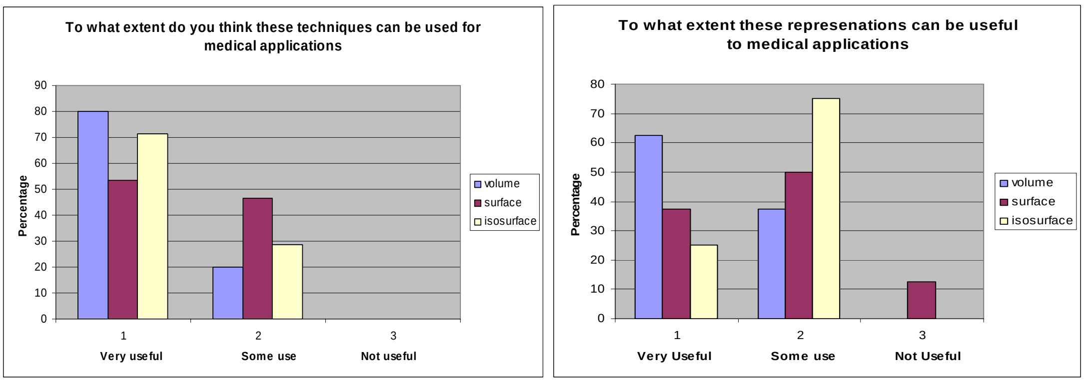


Figure 12 (Group A) & Figure 13 (Group B) – *extent to which different modelling techniques can be used for medical applications*

Figures 14 and 15 explore in more detail the perceptions of the two groups into the effectiveness of the immersive display rather than VR in general. Once again Group A who had the opportunity first hand to experience the immersive nature of the applications respond more favourable to the question, whilst those in Group B were more unsure as to the benefits. It should be noted however that none of the Group B respondents responded negatively to the question.

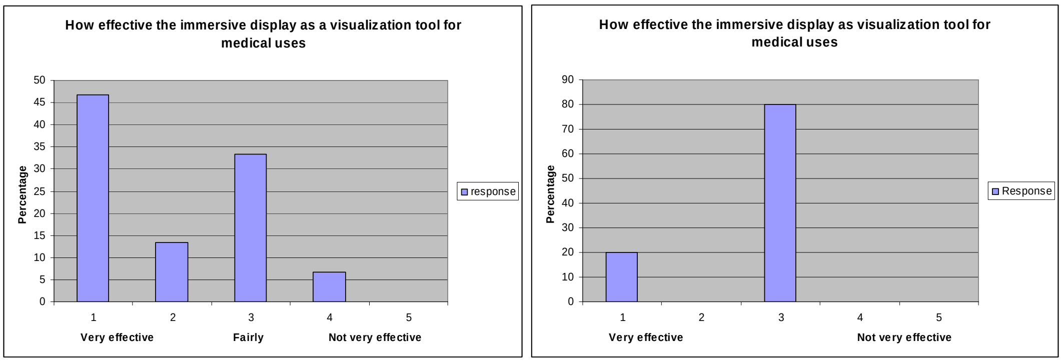
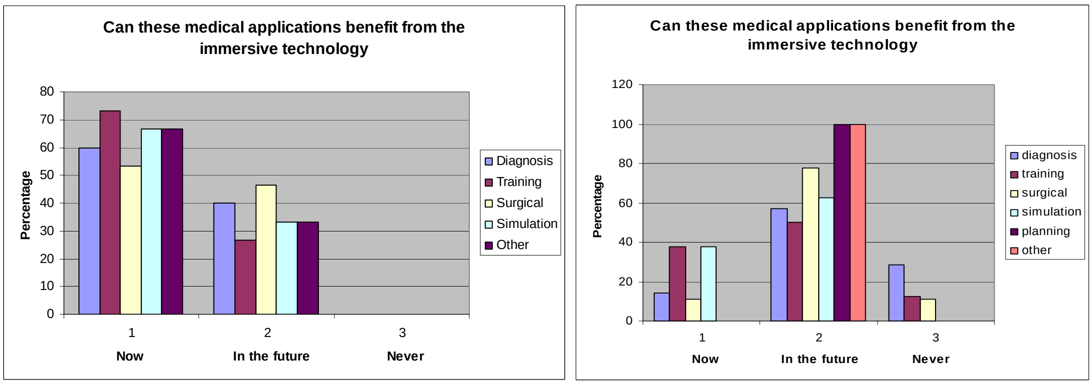


Figure 14 (Group A) & 15 (Group B) – *the effectiveness of immersive display technologies as visualisation tools.*

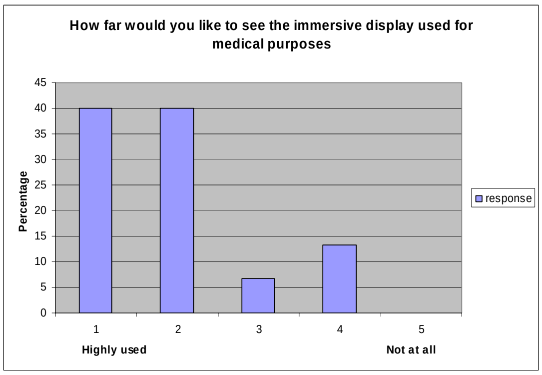
Figures 16 and 17 look in more detail at the type of applications (diagnosis, training, surgical procedures, simulation etc.) the participants believed could most benefit from immersive technology.



Figures 16 (Group A) & Figure 17 (Group B) – *what applications can benefit from immersive technology?*

Group A participants on the whole see the application of the technology to medicine as being much more in the present than those in Group B who see it as a technology for the future. Only a small number of Group B participants and none of Group A participants believe that there will never be any benefit from using immersive technology for medical applications. There seems to be some scope for immersive VR to be used for the range of medical applications – diagnosis, training, surgical procedures, simulation, planning and other non-specified applications.

Group A were also asked about the degree to which they would like to see immersive displays used for medical purposes; how effective they saw immersive displays as visualisation tools and for medical simulations; and whether volumetric representations could help in examining medical structures and identifying disease. Results from these questions were generally favourable and are shown in Figures 18 – 21.



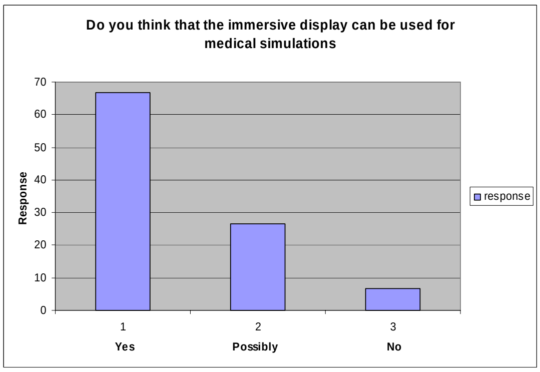
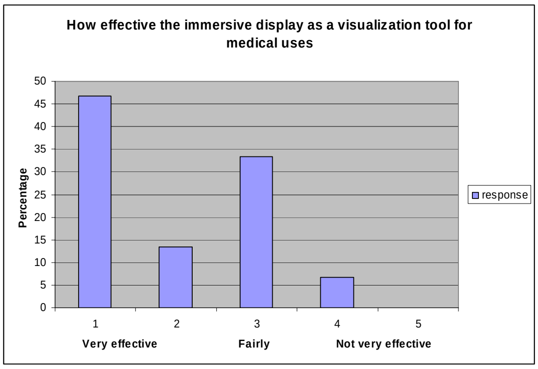
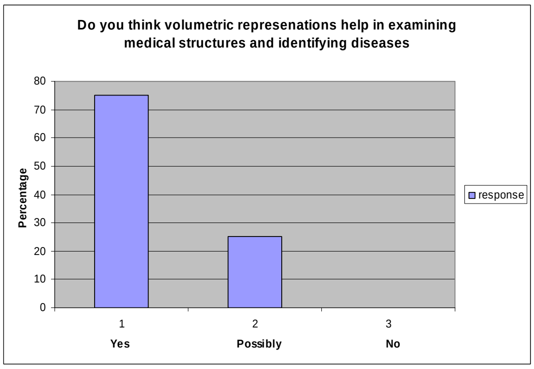


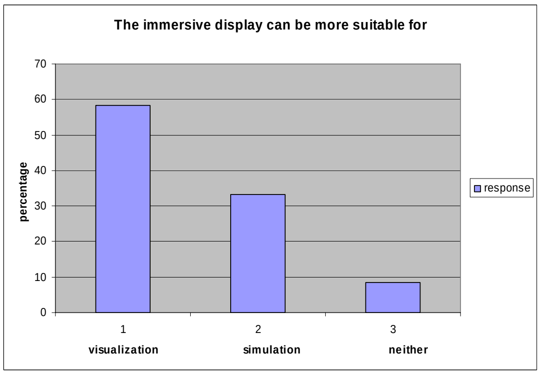
Figure 18 & Figure 19. *The desire to use immersive displays for medical purposes; the use of the immersive display device as a simulation tool (right)*

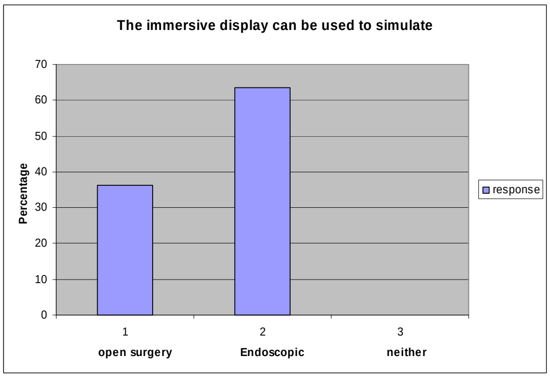




Figures 20 & 21. *The immersive display as a visualisation tool for medical uses; use of the volumetric modelling in examining structures and identifying disease.*

In Group B respondents saw the immersive display as being more useful for visualisation than simulation purposes and of these a greater proportion favoured use for simulating endoscopic procedures than for open surgery simulations as shown in Figures 22 and 23.





Figures 20 & 21. *The immersive display as a visualization tool in medicine; the immersive display device as a simulating tool for surgery (left).*

6. CONCLUSIONS AND FURTHER WORK

In this paper a number of the immersive VR applications that we have developed during the past 18 months using polygonal, iso-surface and volumetric modelling approaches have been described. These models have been demonstrated to medical practitioners who have had the opportunity to interact with them in an immersive environment and assess their usefulness to the medical profession for diagnostic, training and educational purposes. A questionnaire was also sent to other medical practitioners who did not have the opportunity for experiencing first-hand interaction with the models. Initial findings of the survey are encouraging. However, it is interesting to note that those practitioners who received demonstrations of the applications see VR as a current technology whilst those who have not directly experienced its use see it as having potential but as a future rather than current technology. The diversity of expertise in these groups has provided us with rich and valuable feedback on the application of VR in general and immersive technology in particular for various medical applications including modelling for education, visualization for diagnosis and simulation for training. Further work is currently underway with regard to ascertaining students perceptions of VR as a teaching medium.

Acknowledgements: We would like to thank all the participants who took part in our study.

7. REFERENCES

A Al-khalifah, D J Roberts (2004), A Survey of Modelling Approaches for Medical Simulators, *Proc.Intl. Conf. Disability, Virtual Reality and Assoc. Tech.* Oxford, pp. 231-239.

C Cruz-Neira, D Sandin and T DeFanti (1993), Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE, *Proc. of SIGGRAPH '93*, ACM, pp.135-142.

P Gingins, P Beylot, P Kalra, N M Thalmann, W Maurel, D Thalmann and J Fasel (1996), Modeling Using the Visible Human Dataset, *Proc. Medical Informatics Europe, IOS Press*, pp.739-743.

S Gibson, C Fyock, E Grimson, T Kanade, R Kikinis, H Lauer, N McKenzie, A Mor, S Nakajima,

H Ohkami, R Osborne, J Samosky, and A Sawada (1998), Volumetric object modeling for surgical simulation, *Medical Image Analysis*, **2**, 2, pp. 121-132.

P Kalra, P Beylot, P Gingins, N M Thalmann, P Volino, P Hoffmeyer, J Fasel and F Terrier (1995) , Topological Modeling Of Human Anatomy Using Medical Data, *Proc. Computer Animation '95, IEEE Computer Society Press*, pp. 172-180.

A Kaufman, D Cohen and R Yagel (1993), Volume Graphics, *IEEE Computer*, **26**, 7, pp.51-64.

A Kaufman, R Yagel and D Cohen (1993) Modeling in Volume Graphics, Modeling in Computer Graphics, B. Falcidieno and T. L. Kunii, Eds. *Springer-Verlag*, pp. 441-454.

A Kaufman, F Dachille IX, B Chen, I Bitter, K Kreeger, N Zhang, and Q Tang (2000), Real-time volume rendering, *Intl. J. Image Systems and Technology*, 11, pp. 44-52.

W E Lorensen, R Kikinis, S Napel, A E Kaufman and J Flynn (1996), Surface Rendering versus Volume Rendering in Medical Imaging, *Proc. IEEE Visualization*, pp. 439-440.

NLM (2006) US National Library of Medicine, National Institute of Health, Visible Human Project, http://www.nlm.nih.gov/research/visible/visible_human.html.

G Webb, D J Roberts and P M Sharkey (2003), Virtual Reality and Interactive 3D as effective tools for Medical Training, *Proc. Medicine Meets Virtual Reality 11*.

Virtual reality for interactive binocular treatment of amblyopia

P E Waddingham[1,2], S V Cobb[1] R M Eastgate[1] and R M Gregson[2]

[1]VIRART, School of M3, University of Nottingham,
University Park, Nottingham, Nottinghamshire, UK

[2,]Directorate of Ophthalmology, Nottingham University Hospital NHS Trust Queens Campus,
Derby Road, Nottingham, Nottinghamshire, UK

[1]epxpes@nottingham.ac.uk, sue.cobb@nottingham.ac.uk,
richard.eastgate@nottingham.ac.uk, richard.gregson@nuh.nhs.uk

www.virart.nott.ac.uk/ibit

ABSTRACT

Amblyopia, or 'lazy eye', is currently treated by wearing an adhesive patch over the non-amblyopic eye for several hours per day, over a period of many months. Non-compliance with patch wearing is a significant problem. Our multi-displinary team involved clinicians and technologists to investigate the application of VR technology in a novel way. We devised a binocular treatment system in which children watch a video clip of a cartoon on a virtual TV screen, followed by playing an interactive computer game to improve their vision. So far the system has been used to treat 39 children of which 87% have shown some improvement in vision. Vision improvement tended to occur within the first 3-4 treatment sessions. This paper describes research development of the I-BiT™ system. We present a summary of results from clinical case studies conducted to date and discuss the implications of these findings with regard to future treatment of amblyopia.

1. WHAT IS AMBLYOPIA?

Amblyopia, or 'lazy eye', is reduced corrected visual acuity which exists in the absence of any detectable organic disease. Amblyopia can be the result of a squint (strabismic amblyopia), in which both eyes are not straight, a difference of the refractive state of each eye (anisometropic amblyopia) or the result of both squint and refractive inequality (mixed amblyopia, combined anisometropic and strabismic amblyopia), in which there is a squint as well as a stronger corrective glasses lens for one eye. This condition affects 2-5% of the population (Hillis, 1986; Thompson et al., 1991) and is currently treated by wearing an adhesive patch over the non-amblyopic eye for several hours per day, over a period of many months (Cleary, 2000). Although this form of occlusion therapy is successful, success rates are variable (Hiscox et al., 1992) and non-compliance with patch wearing is a problem that can result in unsuccessful treatment (Louden, Polling and Simonsz, 2002).

Standard teaching has been that amblyopia caused by strabismus and anisometropia should be treated before a child is 7 years of age (von Nooden and Crawford, 1979) and research studies have found that screening at a younger age (i.e. under 3 years) leads to better treatment outcomes (Williams et al., 2003). Children who are not successfully treated within this critical period will be left to cope with their condition into adulthood. Whilst, not disabling in itself, having one weak eye can prohibit people from some occupations such as those driving professions which require a HGV licence, the police force and ambulance driver, and there is also a higher risk of losing the good eye in later life due to injury or eye disease. The UK visually impaired register consists of large numbers of people who are amblyopic, but have lost their remaining 'good'eye through injury or disease (Rahi et al., 2002; Tommila and Tarkkanen, 1981).

2. INTERACTIVE BINOCULAR TREATMENT

A research collaboration between orthoptists, ophthalmologists and the virtual reality applications research team (VIRART) examined the potential application of VR technology to treatment of amblyopia. Our multi-displinary team considered how the features of VR could be used to provide a new way to treat amblyopia that young children would find interesting and so encourage them to comply with treatment. Together we devised a novel system providing interactive binocular treatment for amblyopia; the I-BiT™ system (Eastgate et al., 2006).

The basis of this new approach for amblyopia treatment is preferential stimulation of the amblyopic eye, achieved by presenting separate (but visually related) images, one to each eye independently. It was initially considered that this treatment may be more effective than patching occlusion merely due to treatment compliance; our expectation was that young children would find it more attractive to watch cartoons and play computer games than wearing an eye patch and so they would be prepared to follow a course of treatment. At the project outset, we had no indication as to how much treatment would be needed. Patching occlusion therapy can take up to 400 hours of treatment in total (Cleary, 2000).

One assumes that prolonged patching treatment is required as time is required for cell growth and 'neural wiring' of receptor cells for the amblyopic eye. However, the neurology underlying the human conditions of strabismus and amblyopia remains elusive (Barrett et al., 2004). If vision is blurred or one eye is covered during the critical period in postnatal development, neurons in the visual cortex lose their responses to stimulation through that eye within a few days. Anatomical changes in the nerve terminals that provide input to the visual cortex have previously been observed only after weeks of deprivation, suggesting that synapses become physiologically ineffective before the branches on which they sit are withdrawn. (Antonini and Stryker, 1999). So if adequate stimulation is lacking during a critical or sensitive period in early childhood, certain cortical functions such as sight will never develop properly later on (Sengpiel, 2005).

Furthermore, we were uncertain regarding which conditions I-BiT™ treatment would be most suitable for. We were not even sure that children with very poor vision in the amblyopic eye would be able to see what was on the computer screen at all as the brain tends to suppress or disregard the content seen by the amblyopic eye. Standard orthoptic tests demonstrate suppression in most children with dense amblyopia (very poor vision) (Waddingham. Personal communication).

3. WHY VIRTUAL REALITY?

Virtual Reality systems comprise computer generated virtual environments which may be representations of real world environments, simulated or abstract environments (e.g. Ellis, 1994). When using a VR system, a stereo image is perceived as a result of viewing two images of the same scene, one presented to each eye, at slightly offset viewing angles which correspond to the different viewpoints our left and right eyes get when viewing the real world. This stereo effect is intended to produce a sensation of depth; of seeing the virtual environment in 3D and therefore perceiving it as a "virtual world". This sensation is further enhanced by viewing the virtual environment via an immersive display such as a head-mounted display (HMD) or CAVE system.

It was this aspect of VR that was important for amblyopia treatment. The basis of the I-BiT™ system was that we could present an image separately to each eye. Not for the purpose of providing a stereo image, but as a means of presenting different visual content to each eye, in which at least one of the images includes dynamic stimuli. This is illustrated in Figure 1., showing a fish tank constructed in an empty Superscape® 3D environment. The basic principle is that the amblyopic or lazy eye is shown the interesting bit (the fish), whilst the good eye is shown the less interesting bit (the background). When viewed through a stereo viewer the patient should see the fish in the tank. Therefore, the amblyopic eye received preferential visual information. To make sure that the patient can fuse the image (line the two images up correctly) there are a significant number of elements (the base of the tank and the plants) common to both images.

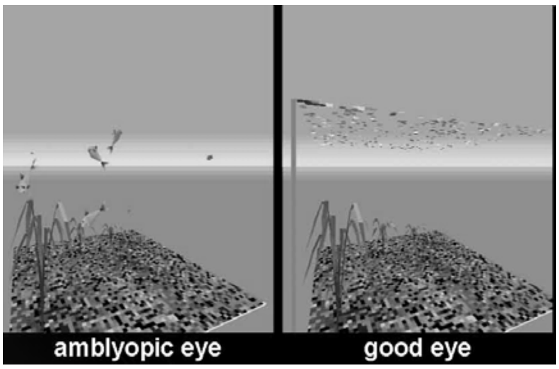


Figure 1. *Different visual information is presented to each eye.*

It was considered that VR could be used for I-BiT™ presentation because it offers;

- Control over the content, resolution and contrast of images presented to each eye
- Control over the visual angle to compensate for any deviation of the eyes and IPD

4. SELECTING A SUITABLE DISPLAY DEVICE

A desktop computer alone does not provide a 3D experience unless attached to a head mounted display (HMD) or an alternative display to provide a 3D effect. We needed to find a 3D viewer that would be suitable for children in the range 5 to 8 years of age. HMDs such as the Virtual Research V8 were too heavy for extended use by our expected user population and children found the headset too claustrophobic. The centre of gravity is at the front which pulls the headset forward, causing neck strain. In addition the inter-pupillary distance (IPD the distance between the two pupils) did not go down to small enough levels for children.

The Cy-Visor DH-4400 binocular headset, although not designed specifically for children, is much more lightweight than the V8, it has a less enclosed design the IPD was small enough for the children we tested. However, although the Cy-Visor could be worn for 30 minutes of treatment, it needed to be held in place by the clinician whilst the child was slightly reclined on the seat.

We decided that we needed to develop our own display in order to conduct preliminary case study trials. The system we developed incorporated a device called a Cyberscope (Wired, 1993). A cyberscope is a plastic hood which fits onto the front of the 15" PC monitor to allow it to be used as a stereo display. This device uses a combination of mirrors to take the image from a standard monitor and present one half of it to one eye and one to the other eye. The monitor display is divided vertically down the middle, and in each half the image to one of the eyes is displayed rotated through 90º to give the images the desired proportions. The mirrors in the Cyberscope rotate the images back the right way up and send them, one to each eye (as illustrated in Figure 2). In this example, both eyes are presented with an image of the clock surround. The left eye is also presented the clock numbers whilst the right eye is presented with the clock hands. Only by using each eye can the patient see what time is shown on the clock.

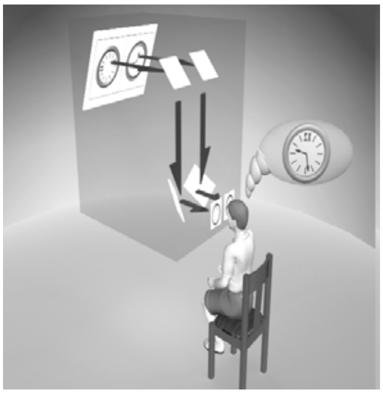


Figure 2. *Concept of the desk mounted display, showing different images presented to each eye.*

By using an LCD monitor we could make the entire display system small and light enough to be fitted into one box. The box was designed such that, when placed on a standard desktop, and used in conjunction with an adjustable chair, the viewing holes are at a height suitable for children.

As the visual scene was visible only to the patient, we needed to provide a secondary screen display for the clinician so that they could see what the patient was doing and check that they were attending to the display with both eyes. This secondary display was also used as the practitioner's control interface, allowing clinicians to select content of the patient's display and configure software adjustments for:

- the viewing angle to each eye
- calibration of IPD
- selecting which eye is the amblyopic eye and which is the non-amblyopic eye
- These controls are described in more detail in (Eastgate et al., 2006).

Figure 3 shows the research prototype system in use.

Figure 3. *Desk mounted version, showing child looking through cyberscope and clinician PC.*

5. SOFTWARE DEVELOPMENT

The virtual environment needed to be capable of presenting less detail to the non-lazy eye and more to the lazy one. Thus, two versions of each virtual environment were built; one containing visually rich and dynamic objects and the other containing static and less interesting objects.

Two games were developed for the first version of the system. A virtual maze game based on the popular 1970's pac-man arcade game (see Figure 4) and a racing game (Figure 5). Pac-man appeared to be ideal for the I-BiT™ system as the game content is comprised of different components: a central character that must navigate around a maze avoiding the "ghosts", whilst attempting to collect all the yellow "dots". The first decision to be made was regarding the level of detail to be displayed in each eye. Essentially there are two distinct groups of components in the game; objects that move and objects that do not. As a result of this, the design decision was to have one eye see the pac-man character and ghosts, whilst the other saw the maze and yellow dots. Also, in order to ensure that the difference was not too great, some continuity between the two was incorporated, in this instance, some elements of the maze were included within the image containing the moveable objects.

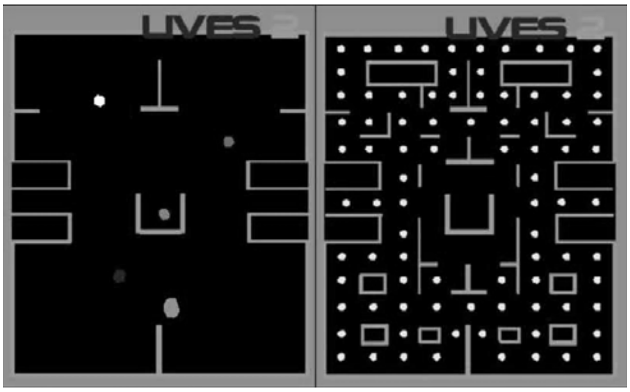


Figure 4. *The virtual pac-man game.*

Although pac-man seemed an ideal game choice, offering a simple game to construct that is easy to understand and gender non-specific, in our trial studies children did not find the game exciting and interesting to play. They complained of limited motion, limited routes around the course, too many ghosts moving in all different directions. It was easy to get trapped and then eliminated by the 'ghosts', which meant that the game wins more often than the child. There was no facility to record performance timings or achieve a score and so the children were not motivated to continue playing the game.

Figure 5. shows the racing game (version 2 with additional icons). In order to get the child to focus on the game, each eye was presented with half a target and alternating white lines in the centre of the road. This

meant that one eye sees a white line, whilst the other eye sees the next white line, therefore combined, it looks like a continuous flow. Some features of the track were presented to both eyes and features of the surrounding environment were presented to each eye separately. The patient had to drive around the track and try to collect the target icons. At the end of the game the number of icons collected and laptime were displayed. This encouraged children to continue playing the game to improve their score. An equal number of icons were presented to each eye and the performance score from each eye was displayed on the clinician's screen. This enabled the clinician to determine whether the amblyopic eye was being used.

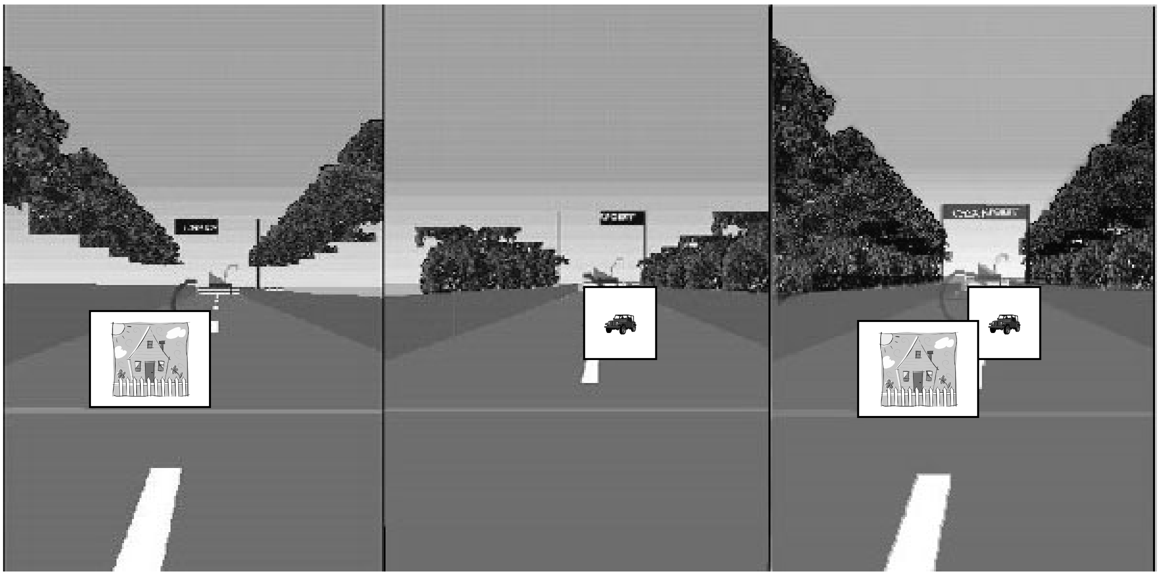


Figure 5. *The racing game as presented to the left eye, the right and the composite view of both images.*

These games were not sufficient to maintain the childrens' interest for continuous play of more than 10 minutes. As we required 30 minutes visual treatment for each session, the system was adapted such that we could present video images via the virtual environment. A virtual TV screen was constructed and video content was presented to the amblyopic eye only. The periphery of the TV surround was presented to both eyes to allow fusion and binocular viewing to occur. A selection of 20-min video clips were available for the children to choose from.

6. CLINICAL CASE STUDIES

We have conducted a number of studies to assess suitability of the I-BiT™ system as a viable treatment of amblyopia. These studies have examined different aspects system design and use. Table 1 presents a summary of clinical case studies conducted to date. These were conducted at two sites by two clinical research teams; the orthoptic unit at Queens Medical Centre, Nottingham and the orthoptic unit at Gartnaval Hospital, Glasgow. Studies are represented using a number system as follows:

- [1] Usability and acceptance of the system by patients, parents and clinicians
- [2] Nottingham studies:
  - [2a] Pilot study, proof of concept using cyberscope desktop viewer, assessment of I-BiT™ system as a treatment for those children who were occlusion treatment failures and children who required occlusion treatment but refused to start the therapy (Waddingham et al., 2006)
  - [2b] Pilot study, single case study using Cy-Visor headset
  - [2c] Follow-on study, effectiveness of the I-BiT™ system as a primary treatment using the desk mounted version (Waddingham et al., in preparation)
- [3] Glasgow studies:
  - [3a] Single case study, same experimental design as study 2b, with the addition of VEP measurement
  - [3b] Follow-on study, effectiveness of using Cy-Visor system as a secondary treatment (Cleary et al., submitted)

Patient composition and trial conditions for each study are listed in Table 1. Performance measures varied according to the study requirements. The usability study consisted of a questionnaire-based assessment for

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

both child and parent and demonstration of the system to children and clinicians. Feedback from this study fed into the pilot studies. In studies 2a, 2b, 3a and 3b visual acuity (VA) was measured before and after each treatment session using LogMAR Glasgow Acuity Cards (McGraw, Winn 1993). In studies 2c and 2d VA was only measured at the end of the treatment session as examination of the data from the previous studies did not show any great variation in vision pre and post session and added another 10 minutes to the session time. The Glasgow Acuity Cards measure visual acuity in consistent gradation from 6/38 to 6/3 at 3 metres testing distance. The patient is shown a flip chart card, which has 4 letters per line. Debate exists to what is considered a significant improvement in vision but for logMAR charts it has been suggested that a change of 0.200 log units (2 lines or 8 letters on the vision test we used) should be used to ensure a real change has occurred (Stewart, 2004) For this reason, change in visual acuity is also presented in terms of percentage proportional improvement (Stewart et al 2003). This measures the changes in vision from the start of treatment to the end of treatment for each eye and is more meaningful to patients (and parents). Visual Evoked Potentials (VEPs) is a method of measuring the electrical responses of the visual part of the brain to visual stimulation and is applied as an objective measure of vision.

Table 1. *Summary of clinical case studies using I-BiT™ system for amblyopia treatment.*

Study and patient groups	Trial conditions	Results
[1] Usability study (Nottingham). 15 Children waiting for a paediatric clinic were asked to participate. 4 sessions	▪ Cyberscope desk-viewer ▪ Games and video clip	▪ Easy to use and very well liked by children. ▪ More activities are required within the system to maintain children's interest over numerous repeated sessions.
[2a] Preliminary pilot study (Nottingham) to determine effectiveness for amblyopia treatment ▪ 6 children aged between 5-7. ▪ 3 treatment failures, 3 treatment refusers	▪ Cyberscope desk-viewer ▪ 2-3 sessions per week for 6 weeks ▪ 20 min cartoons + ▪ 10 min VR game (choice of 2 games)	▪ Significant improvement in VA (average increase 13 letters on LogMAR Glasgow cards) in 5/6 children. ▪ Mean increase = 10 letters ▪ Plateau at session 8 ▪ 3/5 children have maintained or improved on the final VA (11-19 months follow up). 1/5 had atropine with further significant improvement. ▪ Ave. 42% proportional improvement (range 14%-92%)
[2b] Single case evaluation to compare headset with desk-viewer. One 9-year old girl	▪ Cy-Visor Headset ▪ 4 times a week for 2 weeks (8 sessions)	▪ Dramatic improvement of 13 letters. 8 letters after session 2 ▪ Ave. 32% proportional improvement
[3a] Single case evaluation to compare VA results with VEPs. (Glasgow)	▪ Cy-Visor Headset Daily session for 5 days: ▪ 30 min I-BiT + ▪ 40 min VEP	▪ VEP results did corroborate VA. ▪ Study not well tolerated by patient as each session took 1.5 hours to complete ▪ VA increased up to session 4
[3b] Treatment study (Glasgow) to determine effectiveness as a secondary treatment of amblyopia ▪ 12 children aged 6-10 ▪ All treatment failures or refusers	▪ Cy-Visor Headset Weekly sessions over 12 weeks: ▪ 20 min I-BiT video ▪ 5 min I-BiT game	▪ Significant improvement in VA in 11/12 children ▪ Improvement ranged 4-15 letters ▪ 2 children VA came up to 6/6 ▪ Significant improvement occurred within 2-3 sessions of treatment ▪ Ave. 35% proportional improvement
[2c] Treatment study (Nottingham) to determine effectiveness as a primary treatment of amblyopia ▪ 19 children aged 4-10 years ▪ No previous treatment for amblyopia	▪ Cyberscope desk-viewer Weekly sessions over 12 weeks: ▪ 25 min I-BiT video ▪ 5 min I-BiT game ▪ 40 min VEP (at visit 1, 6 & 12)	▪ 17/19 improved vision ▪ Improvement range 2-20 letters (average 7.8 letters) ▪ Most improvement within first 4 sessions ▪ No correlation VEP with VA ▪ Poor patient tolerance of VEP ▪ Ave. 34% proportional improvement (range 11%-67%)

7. RESULTS

Study [1] assessed usability and acceptance of the system by patients (i.e. children), parents and clinicians. The outcome of this study showed high levels of interest in using the system. Patients preferred this option to receiving occlusion treatment at home and parents said that they would be prepared to have regular visits to the clinic for their children to receive I-BiT™ treatment. Clinicians were also positive about the concept of I-BiT™, although expressed some concerns over the need for additional clinic visits for patients, increasing clinician workload and patient waiting times.

The Nottingham studies [2] were all conducted using the cyberscope desktop viewer. The first pilot study conducted to demonstrate proof of concept [2a], yielded extremely positive results showing exceptional improvement in vision for children with very dense amblyopia. This surprising result demonstrated that children with very poor vision could see and respond to the visual display with their amblyopic eye. Moreover, the significant improvement in vision over a relatively short period of time (6 weeks) compared to traditional occlusion therapy (several months), suggests that the neurology of amblyopia is not time-dependent. The follow-on study [2b] provided further indication of this as dramatic improvement in VA was observed after only two sessions (one hour of total treatment). However, although vision improved in 17/19 patients, the degree of improvement was variable, suggesting that children respond differently to the treatment as in traditional patching therapy. There was no pattern of response to the different types of amblyopia. This could suggest that different types of amblyopia exist, rather that the traditional classification currently used. In two of the case studies VEP measurement was applied as an objective means of verifying the subjective VA scores. However, although the first study [3a] of one patient showed a correlation between the increase in vision and the response in the brain, this was not replicated in the Nottingham study [2c] and the method of tested was poorly tolerated by the four children who had it carried out.

The Glasgow studies [3] were all conducted using the Cy-Visor headset. The pilot case study [3b] verified acceptance and reliability of the Cy-Visor headset as an alternative display, despite the need for a reclined patient posture and adult supervision required to hold the headset in place. VEPs measures were never implemented in the Glasgow studies [3b] due to availability of equipment. The Glasgow treatment study corroborated the findings of study [2a] with a larger sample of occlusion treatment failures and/or refusers. Statistical analysis of change in VA, tracked over time, verified the rapid effectiveness of I-BiT™ treatment: No significant change in VA occurred after three sessions (Waddingham et al, in preparation) or four sessions (Cleary et al., in preparation) indicating that the treatment effect occurs within two hours of total treatment time or less.

Despite differences between these studies having used different display devices, in studies applied by different research teams at different clinic locations, the overall pattern of results are similar. Examination of the proportional improvement figures shows an average of 32%-42% across all studies, with a range of 11%-92% for individual patients. The variability in effectiveness may be due to individual differences in patient conditions individual circumstances: the children had a range of different causes of their amblyopia (squint and anisometropia), a wide range of initial vision (6/15 - 6/120) and, at 5-10 years of age were older than generally considered ideal for occlusion treatment. Some of these patients had previously failed to improve with occlusion treatment, yet displayed improvement in vision using the I-BiT™ system.

8. CONCLUSIONS

We have developed a demonstrator that shows how VR technology can be used for binocular interactive treatment of amblyopia (lazy eye). Our clinical case studies provide encouraging results. Patients were keen to use the system (only 2 /39 children did not want to use it) and the rapid improvement in VA from the first treatment session provided immediate positive feedback to children and parents. Moreover, as treatment effectiveness appears to occur in under 2 hours, this method of treatment would not require extensive clinic visits over long periods as initially feared.

Treatment response was variable. For some patients, I-BiT™ treatment alone was sufficient to improve vision to a satisfactory level where no further treatment was required (Cleary et al., submitted). In other patients, residual amblyopia remained and VA was further improved with traditional amblyopia treatment such as occlusion patching (Waddingham et al., 2006). Further work is required to determine when and how I-BiT™ treatment should be applied but on the basis of our case studies, we would suggest that it may be offered as a stand-alone secondary treatment or as a first-step primary treatment to bring VA up to a level from which other amblyopia treatment methods may be continued. It is hoped that an initial improvement in VA will create a positive relationship between patients (parents) and clinicians and thus improve continued treatment compliance.

I-BiT™ treatment worked more quickly than anticipated and not in the pattern expected. This raises some questions regarding the underlying neurology of amblyopia treatment. Rapid increase in VA of the amblyopic eye suggests that there is not enough time for cell growth to take place. We consider that this treatment method is enabling reactivation of dormant neural pathways instead. It follows that, if this is the case, then amblyopia treatment may not be restricted to the critical period of brain development (under 8 years of age) and therefore may be applicable to older children and adults. Certainly, there is interest from the adult amblopia population to further investigate this: in response to a recent news item describing the I-BiT™ system (Mitchell, 2006), we have received over 200 requests from adults seeeking treatment.

This is an exciting research area in its infancy. One of the challenges for further development is design and evaluation of visual content to provide stimulation for maximum treatment effectiveness that is attractive to all patients.

Acknowledgements: The authors would like to thank all members of the I-BiT™ research team at VIRART and Queens Medcial Centre, Nottingham. We also thank the clinical research team at Gartnaval Hospital, Glasgow. We thank all children and parents who were willing to participate in these studies. This work was possible due to funding from the Willaim and Mabel Morris Charitable Trust, the University of Nottingham and University of Nottingham NHS Trust and Carleton Ltd. The I-BiT™ technology has been licensed to Carleton Optical Equipment Ltd. VIRART is part of the INTUITION European Network of Exellence inVirtual Reality.

9. REFERENCES

Antonini A, & Stryker MP, (1999). Rapid remodelling of axonal arbors in the visual cortex, *Science*, 260 1819-1831.

Barrett BT, Bradley A, & McGraw PV (2004). Understanding the neural basis of amblyopia. *The Neuroscientist*, 10 (2), 106-117.

Cleary M. (2000). Efficacy of occlusion for strabismic amblyopia: can an optimal duration be identified? *Br J Ophthalmol*., 84, 572-578.

Cleary, M, Moody, AD, Buchanan, A, Stewart, H & Dutton, GN. (submitted). Assessment of a computer-based treatment for older amblyopes: The Glasgow Pilot Study. *Eye*.

Eastgate RM, Griffiths GD, Waddingham PE, Moody, AD, Butler, TKH, Cobb, SV, Comaish, I, Haworth, SM, Gregson, RM, Ash, IM & Brown, SM (2006). Modified virtual reality technology for treatment of amblyopia. *Eye*, 20(3), 370-374.

Ellis, S.R. (1994). What are virtual environments? *IEEE Computer Graphics ans Applications*, 14 (1), 17-22.

Hillis, A. (1986). Preventable, Curable, Neglected. *Public Health Review*, 14 (3-4), 213-235.

Hiscox F, Strong N, Thompson JR, Minshull C, & Woodruff G. (1992). Occlusion for amblyopia: a comprehensive survey of outcome. *Eye*, 6, 300-304.

Loudon SE, Polling JR, & Simonsz HJ. (2002). A preliminary report of about the relation between visual acuity increase and compliance in patching therapy for amblyopia. *Strabismus*, 10(2), 79-82.

Mitchell, G. (2006). Video games tackle ‘lazy eye’. Digital Planet, BBC World Service. Appeared on *BBC News online*, Tuesday 28th March 2006. http://news.bbc.co.uk/1/hi/technology/4849244.stm

McGraw, PV, & Winn, B. (1993). Glasgow Acuity Cards: A new test for the measurement of letter acuity in children. *Ophthal Physiol Opt*, 13, 400-404.

Rahi JS, Logan, S, Borja, MC, Timms, C, Russell-Eggitt, I & Taylor, D. (2002). Prediction of improved vision in the amblyopic eye after visual loss in the non-amblyopic eye. *The Lancet*, 360;9333, 621-622

Sengpiel, F. (2005). Visual Cortex: Overcoming a No-Go for Plasticity. *Current Biology*, (15), R1000-R1002

Stewart, C. (2004). LogMAR vs. Snellen visual acuity measurements: improved accuracy for the detection if subnormal vision and change, *Eye News* Aug/Sept, 33-36.

Stewart, C. et al (2003) Defining and measuring treatment outcome in unilateral amblyopia. *British Journal of Ophthalmology*, 8, 1229-123?

Thompson JR, Woodruff G, Hiscox FA, Strong N, & Minshull C. (1991). The incidence and prevalence of amblyopia detected in childhood. *Public Health*, 105, 455-462.

Tommila V, & Tarkkanen A, (1981). Incidence of loss of vision in the healthy eye in amblyopia. *Br J Ophthalmol*, 65, 575-577.

von Nooden, GK & Crawford, ML., (1979). The sensitive period, *Trans Ophthalmol Soc UK*, 99, 442-446.

Waddingham PE, Butler TKH, Cobb SV, Moody, ADR, Comaish, IF, Haworth, SM, Gregson, RM, Ash, IM, Brown, SM, Eastgate, RM & Griffiths, GD, (2006). Preliminary results from the use of the novel Interactive Binocular Treatment (I-BiT™) system, in the treatment of strabismic and anisometropic amblyopia. *Eye*, 20(3), 375-378.

Waddingham, PE. (personal communication). Suppression in children with dense amblyopia: 15 years’ experience of orthoptic examination.

Waddingham, PE (in preparation). The application of VR in the treatment of amblyopia. PhD Thesis, University of Nottingham, UK.

Waddingham, PE, Rotchford, A, Cobb, SV, Herbision, N, Gregson, RM, Ash, IM & Eastgate. RM (in preparation). The use of the novel Interactive Binocular Treatment (I-BiT™) system in the treatment of strabismic and anisometropic amblyopia.

Williams, C., Northstone, K., Harrad, R.A., Sparrow, J.M. & Harvey, I. (2003). Br. J. Ophthalmol. 87, 988-993.

TheraGame – a home based virtual reality rehabilitation system

R Kizony[1], P L Weiss[2] M Shahar[3] and D Rand[4]

Laboratory of Innovations in Rehabilitation Technology
Department of Occupational Therapy, University of Haifa, Mount Carmel, Haifa, Israel

[1]*rachelk@zahav.net.il*, [2]*tamar@research.haifa.ac.il*, [3]*smeir@research.haifa.ac.il*, [4]*drand@univ.haifa.ac.il*

ABSTRACT

The limitations of existing virtual reality (VR) systems in terms of their use for home-based VR therapy led us to develop "TheraGame", a novel video capture VR system. TheraGame operates on a standard PC with a simple webcam. The software is programmed using a Java-based visual interaction system. This system enables a quick and easy definition of virtual objects and their behavior. The user sits in front of the monitor, sees himself and uses his movements to interact with the virtual objects. The objective of this presentation is to present the system, a number of the current applications, and some initial pilot usage results. Results from a study of 12 healthy elderly subjects showed moderate to high levels of enjoyment and usability. These scores were also high as reported by 4 participants with neurological deficits. Some limitations in system functionality were reported by one person with stroke who used TheraGame at home for a period of 2.5 weeks. Overall, TheraGame appears to have considerable potential for home based rehabilitation.

1. INTRODUCTION

Clinicians who work in rehabilitation aim to enhance clients' functional ability as well as their ability to participate in community life. These goals are achieved by intensive intervention aimed at improving motor, cognitive and meta-cognitive abilities. Regrettably, for economic reasons the duration of subacute rehabilitation is getting shorter. Therefore supervised self-training at home is necessary in order to achieve maximal functional independence. In addition, during the last decade, there is growing evidence that neurological patients should receive maintenance therapy after rehabilitation in order to preserve the achievements gained while in hospital as well as to prevent deterioration. Moreover, several studies have shown that improvement of impairments and reducing participation restriction can be achieved by training even in the chronic stage (Liepert et al, 2000; Pang et al, 2006).

Virtual Reality-based therapy has well-known assets including the opportunity for active learning in challenging but safe and ecologically-valid environments, while maintaining strict control over stimulus delivery and measurement, and the capacity to individualize treatment needs, while gradually increasing the complexity of tasks (Rizzo et al, 2002; Rizzo et al 2004) Although the advantages of Virtual Reality (VR) are becoming widely recognized within the clinical community, the rehabilitation team faces a challenge when looking for a home-based, affordable VR system that suits the implementation of therapeutic objectives.

Video capture VR consists of a family of camera-based, motion capture platforms that differ substantially from the HMD and desktop platforms in wider use (Weiss et al, 2005). When using a video-capture VR platform, users stand or sit in a demarcated area viewing a large video screen that displays one of a series of simulated environments. Users see themselves on the screen, in the virtual environment, and their own natural movements entirely direct the progression of the task, i.e., the user's movement is the input.

Currently, the main video capture platforms that have been used for rehabilitation are GestureTek's (formerly know as VividGroup) GX and IREX (Interactive Rehabilitation EXercise) (www.irex.com) systems. For the past three years, the Sony PlayStation II EyeToy applications have begun to be used as an intervention tool as well. The GX-IREX platforms have been adapted for rehabilitation but are considerably more expensive than the EyeToy, and require a more elaborate setup including a chroma key blue/green backdrop behind the user and bright, ambient lighting. Sony's EyeToy is an off-the-shelf, low-cost gaming application that may be run under almost any ambient conditions. However, the EyeToy cannot be graded to suit clients' cognitive and motor impairments especially when those impairments are severe.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

The limitations of both of these systems in terms of their use for home-based VR therapy led us to develop "TheraGame", a novel video capture VR system. TheraGame is an innovative virtual reality system for dynamic, low-cost rehabilitation treatment with tele-medicine capabilities. It is suitable primarily for use in hospitals, local clinics, chronic care facilities and at home. It enables patients to engage in computer simulation-based, gaming activities that improve their motor, cognitive and meta-cognitive abilities while measuring performance. The system is based on low-cost, technically simple components and thus it is accessible to a wide range of consumers in different cultural settings. Furthermore, it can increase the amount and intensity of simple, repetitive tasks that patients can do on their own or with minimal supervision, hence releasing the therapist from constant supervision of rote exercise, and reducing the overall cost of rehabilitation. At the same time, the system provides an accurate outcome measure that assists in tracking the patient's condition and generating individualized programs for each patient. An equally important application of this product is its use by the healthy population of all ages, enabling them exercise and maintain a healthy life style in an enjoyable manner.

TheraGame aims at combining the strengths of both the GX and the EyeToy system – it is low cost (similar to the EyeToy) while maintaining the flexibility and the ability to program and change the system (similar to the GX system).

The objectives of this paper are: 1. to present the system, and a number of the current applications, and 2. to present initial pilot usage results of an ongoing study, with elderly people as well as people with neurological disabilities.

2. METHODS

2.1 Participants

To date, 12 healthy elderly participants, seven female and five male, have experienced TheraGame. Their mean ± standard deviation (SD) age was 70.6 ± 4.4 years and ranged between 65 to 78 years. Ten of these participants reported that they use a computer on a regular basis for a mean 2.2 ± 2.3 hours per week.

A second group included four participants, aged 65-76 years, who experienced TheraGame for one session; three were post stroke and one had had a spinal stenosis. All of the participants had a weak upper extremity; three participants could use both hands while one could use only one hand. In addition, one participant, a 53 year old man who had a stroke two years prior to the study, used a system that was installed on his own computer at home for a period of two and a half weeks. This participant suffered mainly from cognitive deficits such as memory loss and meta-cognitive deficits such as impaired executive functions.

2.2 Instruments

2.2.1 VR system – TheraGame. TheraGame operates on a standard PC with a simple webcam. The software is programmed using a Java-based visual interaction system. This system enables a quick and easy definition of virtual objects and their behavior. The user sits in front of the monitor, sees himself and uses his movements to interact with the virtual objects. As illustrated in Figure 1, the controllers (i.e., virtual buttons or arrows) are displayed in a separate area on the screen which also displays the user. In this way users receive visual feedback regarding when and how long they have touched the controllers. The controllers operate a game which is presented on a larger portion of the same screen. In addition to providing a platform for the development of new games for TheraGame, the system enables embedding of existing Flash games. To keep the interaction simple, the games need to have basic controls of up to six different keys (usually the arrow keys and the space bar). The number, color, size and location of the controllers may be defined in addition to the game parameters such as speed and level of difficulty. Thus, the level of the games may be graded to the client's level. Since the operation of TheraGame is easy and the cost is relatively low, the goal is for clients to purchase TheraGame following the rehabilitation process, to be used at home independently (with periodic supervision) in order to improve motor, cognitive and meta-cognitive deficits.

Examples of games include Touch-tetris (see Figure 2), virtual ball games, different maze applications, and TheraSlide-Show where a set of pictures (family photos or other pictures or photos) are loaded into the system and presented. The user can use the left and right arrow to scan the photos to look at them or arrange them in order. The principle of all of the applications is similar with that of the other video capture VR systems (Weiss et al., 2005) which is to provide enjoyable and motivating tasks that require varying combinations of motor and cognitive skills. The different games encourage clients to use both or one of their arms (instead of using a mouse and a keyboard) while they are engrossed in another task which diverts their attention away from the difficult motor task and encourages the use of cognitive abilities such as memory, planning, and visual scanning. It is anticipated that this will result in more active movement of the impaired limb than would occur during traditional, rote exercises. Outcome measures include the number of points, levels per game and time of engagement. In addition the system registers the date and time the user played the games.

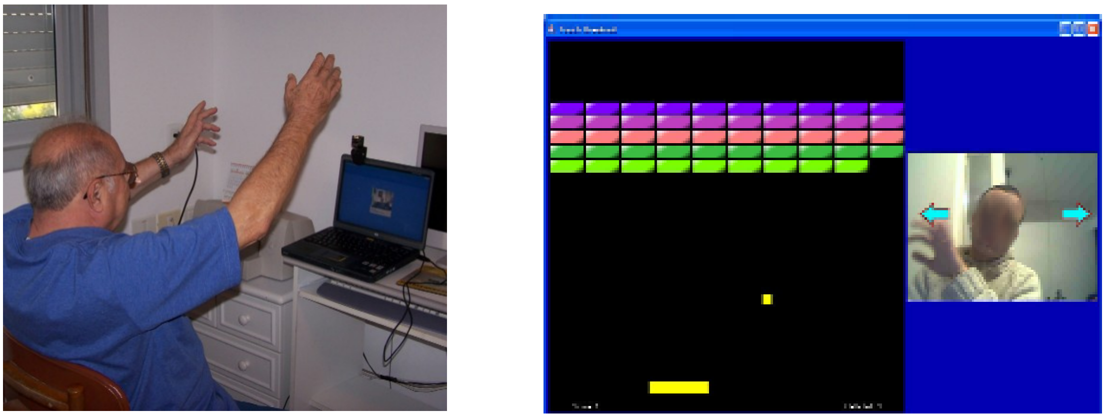


Figure 1. *Inside Motion Tetris (right) and outside (left) a game (Motion Music) in TheraGame.*

The healthy participants experienced two maze like applications which required planning and problem solving: “ColorSok” (see figure 3, right) – where pairs of color blocks must be pushed together and “Frogs” (see figure 3, left), where the user has to help the frog jump on to all of the Lilies in order to make them all disappear.

The stroke participants experienced three differences games including Virtual ball games, Motion music game and Touch-tetris. The patients post-stroke at home experienced Frogs, Touch Tetris, and another two planning games.

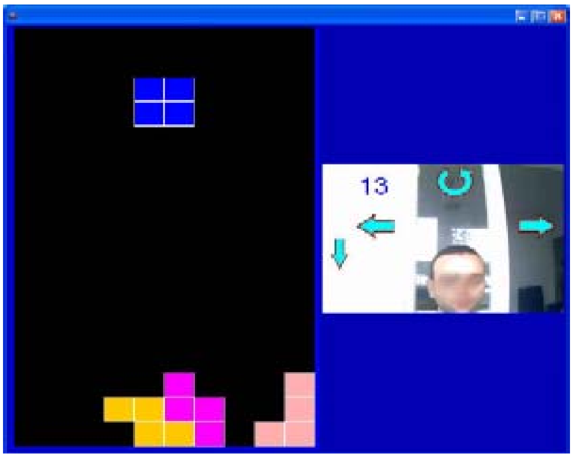


Figure 2. *A screenshot of touch-Tetris in TheraGame.*

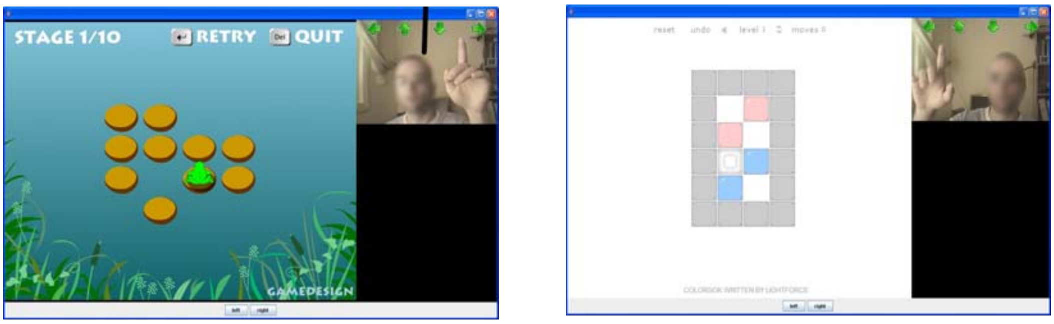


Figure 3. *A screenshot of Frog (left) and ColorSok (right) in TheraGame.*

2.2.2 Short Feedback Questionnaire (SFQ) (Kizony et al, 2005). This is a 7-item questionnaire designed to obtain information about the subjective responses of the participants to the VR experience in each scenario. It queries the user's sense of presence, perceived difficulty of the task and any discomfort that users may have felt during the experience. The first six items of the questionnaire were formulated as an abbreviated alternative to the longer Presence Questionnaire developed by Witmer and Singer's [20]. These items assess the participant's (1) feeling of enjoyment, (2) sense of being in the environment, (3) success, (4) control, (5) perception of the environment as being realistic and (6) whether the feedback from the computer was understandable. The seventh item queried whether participants felt any discomfort during the experience. Responses were rated on a scale of 1-5 where 1 = "not at all" and 5 = "very much". In this paper, responses from questions 1 and 6 are reported.

2.2.3 Borg's Scale of Perceived Exertion (Borg, 1990). This scale was used to assess how much physical effort the participants perceived that they expended during each VR experience. This is a 20-point scale that participants rated from 6 (no exertion at all) to 20 (maximal exertion).

2.2.4 System Usability Scale (SUS) (Brooke, 1986). The SUS includes ten items which provide a global view of subjective assessment of a system's usability. Each item was rated on a 5-point scale from 1 (disagree totally) to 5 (agree totally). Five items are positive statements, such as "I think that I would like to use this system frequently" and "I thought the system was easy to use" and the other five items are negative, for example, "I found the system unnecessarily complex" and "I think that I would need the support of a technical person to be able to use this system". The item scores were totaled as described by Brooke (1986) to give an overall score ranging from 10 to 100 points.

2.3 Procedure

The healthy participants experienced the system for one 20-30 minute session. Each participant was trained to use the virtual arrows within the system using a simple application before the actual experience. After training each participant played two games for 5 minutes each. After each game they completed the SFQ and rated their level of exertion on the Borg scale. To assess the usability of the system the examiner demonstrated how to operate the system (i.e., how to commence a new game); upon completion of the second game the participants were asked to repeat this action, after which they filled in the SUS. The group of 4 participants with neurological deficits also experienced the system, operated by an occupational therapist, for one 30 minute session and played three different games. For the other participant with stroke, a system with four games was installed at his home. Following the installation, an occupational therapist trained the participant and his wife how to operate the system. They were asked to record in a journal when, for how long and what games he played with the system. After a period of two and a half weeks the therapist returned to the participant's home and carried out a structured interview with the couple. The wife also completed the SUS.

3. RESULTS

The reported level of enjoyment was 3.8 ± 1.0 and ranged between 2 and 5. The mean Borg scale for the two games was 11.2 ± 1.8 indicating that the perceived exertion level was low. The mean level of usability according to the SUS was 73.8 ± 14.5 and ranged between 55 and 100, indicating a relatively high level of usability.

The level of enjoyment for the group of 4 participants with neurological deficits was 5. They also reported that the feedback provided by the computer was very clear (SFQ scores of 4-5). Their perceived exertion ranged between 10 (very easy) and 13 (somewhat hard). All participants said they would like to use such a system in their own homes. In addition, they requested an increase the size of the screen and a display of the scores for all the games. On some occasions they had difficulty touching the correct arrow, since their hand touched another arrow by mistake.

The participant who used the system for the extended period of time (2.5 weeks) used it for 10 sessions over 16 days for a total of 213 minutes. He does not suffer from motor impairments but does have memory problems and meta cognitive deficits. His wife's responses to the SUS indicated a high level of usability (score = 82.5/ 100 points). Their responses to the structured interview were variable. When asked "how much did using the system contributed to motor and cognitive training or providing leisure activities?" the participant said that he felt the experience contributed more to motor aspects whereas his wife said it contributed to cognitive aspects such as memory and planning. She also mentioned that it contributed to leisure activities and to rehabilitation in general. Both of them said that it would be better to activate the system with a keyboard and a mouse and not with the virtual arrows. The wife explained that the virtual arrows are distracting. In addition they both said that there is a need to improve the system and to increase the number of games available for the user.

4. SUMMARY AND CONCLUSIONS

The results of this study demonstrate the usability of TheraGame for healthy elderly people and for a small number of people who are post-stroke. This finding is important as a first step prior to the implementation of this type of system in home-based rehabilitation services, especially for people who have had a stroke. Since a large number of stroke survivors suffer from cognitive and motor deficits they could be expected to have difficulty operating such systems on their own, at least initially, and would need help from their spouse who is usually also elderly. The usability of the system with people who have had a stroke or brain injury should be further studied in order to determine the minimum cognitive and motor abilities necessary to operate TheraGame independently.

One of the primary reasons for using VR in rehabilitation is that the gaming factors enhance enjoyment and motivation for treatment (Rizzo and Kim, 2005). Tauer and Harackiewicz (2004) have suggested that intrinsic motivation involves the desire to participate in an activity for its own sake, is marked by high levels of enjoyment during the task and should help the person to stay interested in the activity for a long period of time. They also suggested that task enjoyment influences performance since a person who enjoys what he is doing spends more time developing his skills in a given activity. Bach-y-Rita, et al. (2002) also indicated the importance of a patient's motivation in order to achieve active participation and functional improvement. The results of this study, although they should be interpreted with caution due to the small sample size, showed that the elderly participants reported a moderate level of enjoyment. This finding is similar to the results of a study using games run with the Sony PlayStation II EyeToy system (Rand et al., 2004). The level of enjoyment reported by the participants who had a stroke was higher, and corresponded to the results reported by Rand et al (2004) in the same study. These findings are also in accordance with level of enjoyment reported by people who had a stroke using the GestureTek GX system (Kizony et al., 2004).

This ongoing study is the first step towards showing the considerable potential that TheraGame appears to have for home-based rehabilitation. However, the results are still preliminary and data collection with additional participants who have had a stroke and are still in hospital as well as others who are already at home is currently underway. Future studies will examine the usability of the system with more clinical populations and will study the effects of additional TheraGame applications and improvements to its outcome measures.

Acknowledgements: Third year students of the Department of Occupational Therapy, University of Haifa, for help with data collection of healthy participants.

5. REFERENCES

P Bach-y-Rita, S Wood, R Leder, O Paredes, D Bahr, Wicab, E Bach-y-Rita and N Murrilo (2002), Computer-assisted motivating rehabilitation (CAMR) for institutional. Home and educational late stroke programs, *Topics in Stroke Rehabilitation*, **8**,(4), pp.1-10.

G Borg (1990), Psychophysical scaling with applications in physical work and the perception of exertion. *Scandinavian Journal of Work Health*, **16**, pp. 55-58.

J Brooke (1986), SUS: A quick and dirty usability scale, *In Usability evaluation in industry* (P W Jordan, B Thomas, B A Weerdmeester & I L McClelland, Eds), Taylor & Francis, London, pp. 189-194.

R Kizony L Raz, N Katz, H Weingarden and P L Weiss (2005), Using a video projected VR system for patients with spinal cord injury, *Journal of Rehabilitation Research & Development*, **42**, pp. 595-608.

R Kizony, N Katz and P L Weiss (2004), A model of VR-based intervention in rehabilitation: Relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke, *Proc, Intl. Conf. on Disability, Virtual Reality and Associated Technology*, P Sharkey, CS Lanyi, P Stanton (Eds); September, Oxford, UK.

J Liepert, H Bauder, W H R Miltnse, E Taub and C Weiller (2000), Treatment-induced cortical reorganization after stroke in humans. *Stroke*, **31**, pp. 1210-1216.

M Y Pang, J E Harris and J J Eng (2006), A community –based upper extremity group exercise program improves motor function and performance of functional activities in chronic stroke: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, **87**, pp. 1-9.

D Rand, R Kizony and P L Weiss (2004), VR Rehabilitation for All: Vivid GX versus Sony Playstation II EyeToy. *Proc, Intl. Conf. on Disability, Virtual Reality and Associated Technologies,*P Sharkey, CS Lanyi, P Stanton (Eds), September, Oxford, U.K

A A Rizzo, J C Buckwalter and C Van der Zaag (2002), Virtual environment applications in clinical neuropsychology. In: K. Stanney (Ed*.). The Handbook of Virtual Environments*. Erlbaum Publishing, New York, pp. 1027-1064.

A A Rizzo and G J Kim (2005), A SWOT Analysis of the Field of VR Rehabilitation and Therapy**.** *Presence: Teleoperators and Virtual Environments,* **14**, pp. 119-146.

A A Rizzo, M T Schultheis, K Kerns and C Mateer (2004), Analysis of assets for virtual reality in neuropsychology. *Neuropsychological Rehabilitaion,* **14**, pp. 207-39.

J M Tauer and J M Harackiewicz (2004), The effects of cooperation and competition on intrinsic motivation and perfromance. *J Personal Soc Psycho*, **86**, 849-861.

P L Weiss, D Rand, N Katz and R Kizony (2005), Video capture virtual reality as a flexible and effective rehabilitation tool. *Journal of NeurEngeeniring and Rehabilitaiton,* **12**,1. Available at: www.jneuroengrehab.co /content/1/1/12.

SMART project: application of emerging information and communication technology to home-based rehabilitation for stroke patients

H Zheng[1], R Davies[1], H Zhou[2], J Hammerton[3], S J Mawson[3], P M Ware[3], N D Black[1], C Eccleston[4], H Hu[2], T Stone[4], G A Mountain[3] and N D Harris[4]

[1]School of Computing & Mathematics, University of Ulster, Newtownabbey, N. Ireland, U.K.

[2]Department of Computer Science, University of Essex, Colchester CO4 3SQ, U.K.

[3]Centre for Health and Social Care, Faculty of Health and Wellbeing, Sheffield Hallam University, Howard St, Sheffield S1 1WB, U.K.

[4]School for Health, Royal National Hospital for Rheumatic Diseases, University of Bath, U.K.

[1]*{h.zheng, r.davies, nd.black}@ulster.ac.uk*, [2]*{zhou, hhu}@essex.ac.uk*, [3]*{j.hammerton,s.j.mawson,p.m.ware,g.a.mountain}@shu.ac.uk*, [4]*{hssce, ts241,n.harris}@bath.ac.uk*

ABSTRACT

The SMART project, entitled 'SMART rehabilitation: technological applications for use in the home with stroke patients', is funded under the EQUAL (extend quality of life) initiative of the UK Engineering and Physical Sciences Research Council (EPSRC). The project aims to examine the scope, effectiveness and appropriateness of systems to support home-based rehabilitation for older people and their carers. In this paper, we describe the design and development of a low-cost home-based rehabilitation system. Through the project we have involved end users in the design process and this model can be applied to the design of other healthcare related systems.

1. INTRODUCTION

In the UK, stroke is the most significant cause of adult disability. Reports (DoH, 2000) showed that six months after stroke, 49% of patients need help with bathing, 31% of patients need help with dressing and 33% of patients need help with feeding. Research suggests that intensive and repetitive training may be necessary to modify neural organization (Miltner et al 1999, Rossini et al 2003). However, in the UK, inpatient rehabilitation length of stay for patients with stroke is decreasing, with limited outpatient rehabilitation. Therefore, there is a need to develop a low-cost, accessible, home-based rehabilitation system.

The SMART project, entitled 'SMART rehabilitation: technological applications for use in the home with stroke patients', is funded under the EQUAL (extend quality of life) initiative of the UK Engineering and Physical Sciences Research Council (EPSRC). The project aims to examine the scope, effectiveness and appropriateness of systems to support home-based rehabilitation for older people and their carers (www.shu.ac.uk/research/hsc/smart). The SMART consortium consists of one NHS Trust, four universities and one voluntary organisation, namely Royal National Hospital for Rheumatic Diseases, University of Bath, Sheffield Hallam University, University of Essex, University of Ulster and The Stroke Association.

The rest of this paper is organised as follows. Section 2 introduces the SMART rehabilitation system currently under development. In Section 3, the user involvement design strategy is presented. The motion tracking unit is detailed in Section 4 and the ICT decision platform is presented in Section 5 respectively. In Section 6, outcome measurements are described. Finally, a brief summary of the system usability and some discussion on the SMART system are presented in Section 7.

2. SYSTEM OVERVIEW

The SMART rehabilitation system employs an ambulatory monitoring system linked to an ICT decision support platform that provides therapeutic instruction, supports the rehabilitation process and monitors the

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

effectiveness of rehabilitation interventions upon patient function. Information relating to this process is fed back to patients, their careers or health care professionals.

The system consists of three componnents; (i) motion tracking unit (ii) base station unit (iii) web-server unit (Figure 1).

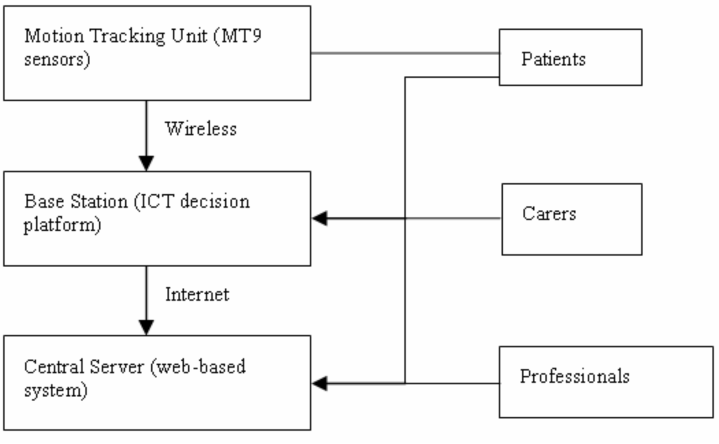


Figure 1. *System Architecture.*

The motion tracking unit (Zhou et al 2004) consists of two MT9 inertial sensors (Xsens Dynamics Technologies, Netherlands) which are attached to the patient's arm to track the movement during activities such as drinking or reaching. The MT9s record the movement information (positions and angles) of three joints, i.e. wrist, elbow and shoulder. The information is then sent wirelessly to the base station (Media PC) via a digital data box called "XBus" (placed on the waist) for further processing by the ICT decision platform (Figure 2). The ICT platform will display the movement in a 3 dimensional (3D) environment at the base station; analyse the data; store the data and upload the data to central server. Healthcare professionals can assess and monitor movements remotely via the internet by accessing the central server, ultimately to provide comments over the web-based system (Zheng et al 2005b). The ICT platform will provide these comments as feedback to the patients and their carers alongside other more detailed analysis.

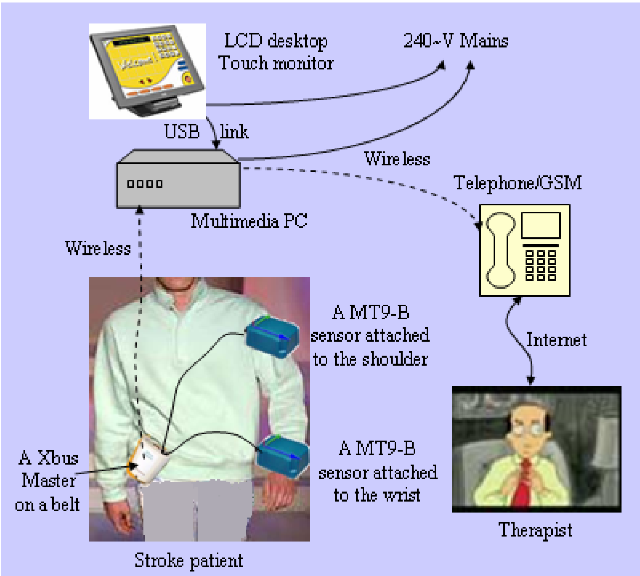


Figure 2. *Illustration of Smart Rehabilitation System.*

3. USER INVOLVEMENT IN DESIGN

Early in the project the project focus groups were held with patients and healthcare professionals to ensure that proposed technical solutions, methodology and outputs were acceptable. A number of key principles were identified with their help during the focus group process:

- It is an aid to therapy, not a stand-alone therapy.
- It is not specific to any one model of therapy
- It is a generic device applicable to a variety of rehabilitation aims for upper and lower limb
- No two people who have had a stroke are the same: there must be flexibility in all elements of the device.
- The device must be as simple as possible to use, and adaptable to individual needs. Stroke patients have complex impairments often incorporating cognitive difficulties such as problems with perception, attention, information processing, language and memory.
- The device should provide accurate feedback on performance.

In the later stages of the project we recruited a group of expert users to provide specific feedback on key aspects of the system such as user interface, type of feedback and computer interface. This data was collected by qualitative researchers, summarised and fed back to the engineering teams. Table 1 summarises some of the key factors that were identified.

Table 1. *Key points for designers: how the individual interacts with the device.*

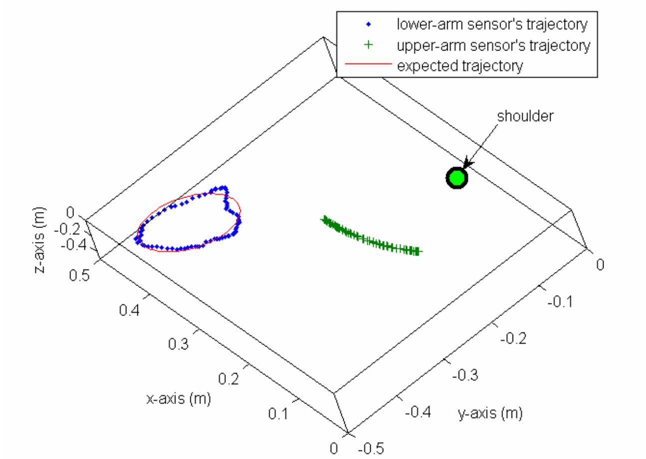
	Implications for design	
Points of interaction		
The sensor (attached to the body)	Ease of application (e.g. one handed, poor fine motor skills, cognitive impairments,) Size of sensor (weight, cumbersomeness) accuracy and repeatability of sensor placement	Independence of use Simplicity of design
The device	Ease of use (e.g. size of buttons, colour codes, information delivery) Instructions for use: (on/off, charging, positioning etc) Capacity to set an individual programme (nb fatigue) comfort and wearability for user adaptability for different users - 'one size fits all'	
The feedback mechanism/s		
Real Time (Knowledge of Performance)	- Choice of methods (auditory, visual, written, storable and retrievable) -Simplicity of information display	Instructions – Different methods / clarity / simplicity Targets, possible to set Accuracy of results
Results for User (Knowledge of Results)	- Choice of methods (auditory, visual, written, storable and retrievable) Feedback presented positively Simplicity of information display	
Results for Therapist (Knowledge of Results)	Visual, written, storable and retrievable records	

4. MOTION TRACKING UNIT

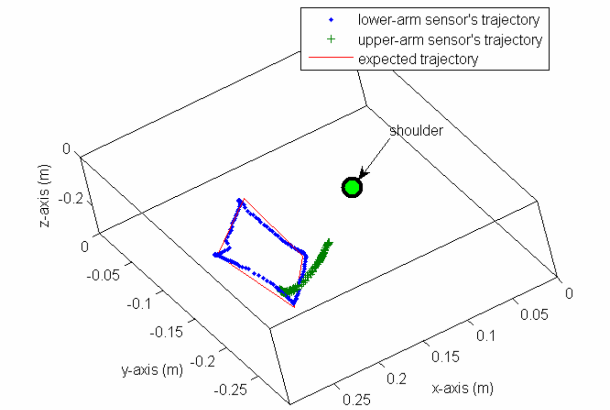
The tracking unit utilises sensor fusion and optimisation techniques and is implemented using Visual C++, based on a Media PC with a VIA Nehemiah/1.2GHz CPU. The Bluetooth wireless feature allows the subject to carry out motion exercises freely.

In order to determine the position of the upper limb, inertial measurements corresponding to human arm movements are continuously generated. A kinematics model is applied to locate the wrist and elbow joints in the global frame. The displacements of the shoulder joint are computed from accelerations of the sensor adjacent to the elbow joint using a Lagrange function with an equality-constrained optimisation method.

The proposed algorithm was validated by tracking a circular motion (radius: 0.1 m) and a square motion (rectangle 0.2×0.14 m^2), drawn on a table. A subject sit still in a chair with the motion patterns placed on the desk in front of him (the lengths of two segments of the arm are 0.26 and 0.24 m respectively). The two MT9 sensors were attached to the middle position of the upper arm and the wrist joint of the lower arm, respectively. During the experiments the subject allowed the MT9 sensor attached to his wrist joint to move along the path of each shape on the desk surface. The data was generated continuously for 40 seconds with a sampling rate of 25 Hz. The errors were defined as the Euclidean distance between the measurements by the MT9 system and the designed trajectories. Means and standard deviations are calculated from these errors (Figure 3).



(a) Mean position estimation of the two sensors in circular motion. Mean errors = 0.017±0.013 m.



(b) Mean position estimation of the two sensors in square motion. Mean errors = 0.011±0.008 m.

Figure 3. *Motion trajectories of the two sensors using the proposed motion detector.*

5. ICT DECISION PLATFORM

The ICT platform consists primarily of five modules: database, sensor interface, decision support, communication and a feedback module.

- The database module stores the personal information, individualised questionnaires to check the patient can safety complete the exercises, rehabilitation history (prior movement data) and the comments/instructions from healthcare professionals;
- The interface module provides tools and menus to access system functions. We have included a facility to allow individual patient to select their own preferences regarding the presentation of the interface, such as colour, font size and feedback style. Users interact with the system using a LCD touch screen monitor.
- The decision support module will analyse the data and provide key outcome variables relating to physical performance (such as length of reach, elbow angle), while the communication module manages the transfer of information with the central server;

- The feedback module is the core module, which provides different types of information to patients, namely 3D movement information, comments/instructions, and analysis results in the following formats: text; 3D visualisation; tabular and graph.

The visualisation feedback displays and replays the movement of rehabilitation exercises to users in a 3D environment. To improve the realism, 3D rendering is applied to a virtual head and arm based on the movement data collected by the tacking unit. In order to provide a reference for patients, stored target movement templates are available which can be overlaid or mirrored on the screen to help the patient replicate the best movement. Figure 4 shows two types of methods used in presenting the 3D information, one displays exercise movement and the target template movement in two separate windows; and the other displays them in the same window with the template movement as a ghost layer. Through preference settings, users are able to choose either mode. This is a novel feature of the interface design, which provides a visualisation that is easy for users to understand, rather than biomechanical stick diagrams. The target movement template is personalised and adaptive. The first target template can be the patient's first movement, after that, it could be the best past movement that the patient has carried out. The decision on which template to use will be made by the therapist based on the change in outcome measurements.

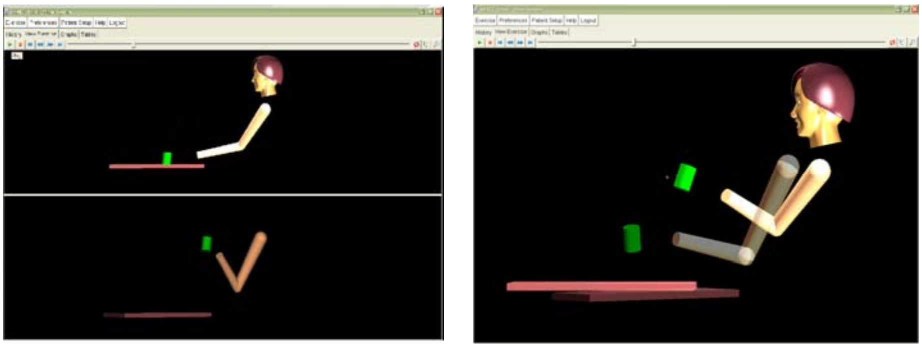


Figure 4. *Screen shot of 3D rendering.*

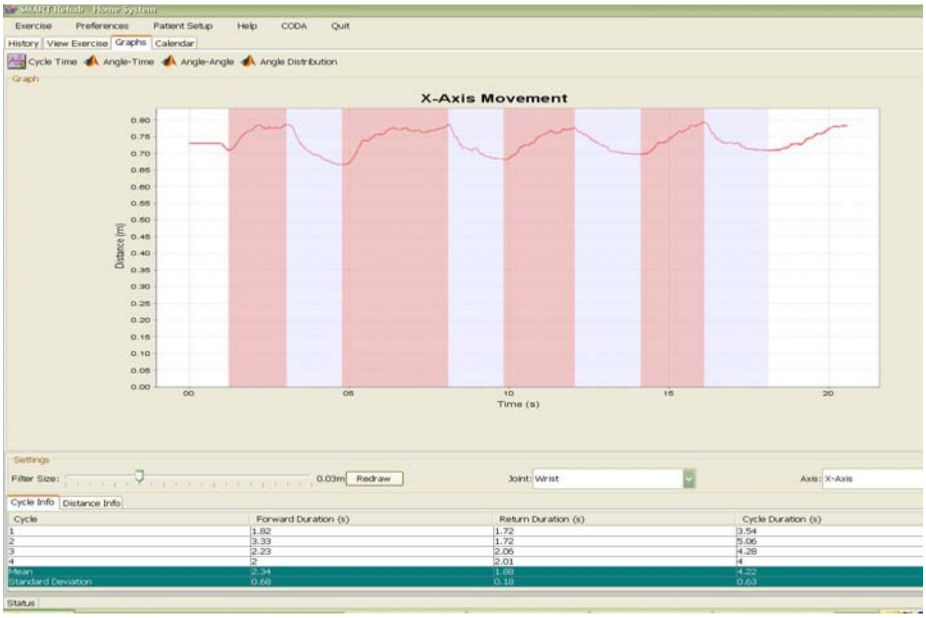


Figure 5. *Screen shot of cycle time measurement*

Measurements of variables are shown in graphs. Figure 5 shows the measurement of cycle time for a reaching movement. The forward duration and return duration of each cycle is calculated and shown. Users can select to view the movement and measurement results of each joint. Mean and standard deviation is calculated for each exercise. Similarly, rotation rate and other variables can be measured and displayed using graphs and/or tables. These variables will be used to monitor and assess the rehabilitation procedure, and can be used to modify the rehabilitation setting and system setting, such as the selection of target exercise template.

 ISBN 07 049 98 65 3

6. OUTCOME MEASUREMENTS

It is important that the system provides outcomes that are clinically relevant to the restoration of functional activities. A range of measurements were identified by the therapy user group and quantified by comparing age matched normative data to stroke data, collected using a commercially available 3D video motion analysis system (CODA Charnwood Dynamics, Rothley, UK).

Work is underway to validate the performance of the MT9 motion sensors against this system prior to evaluation in a home environment and to optimise the user interface following feedback from the expert user group and naive uses attending local Stroke Clubs. We are also evaluating a range of quantitative outcome measures that might be used to provide feedback on the progress of the rehabilitation.

7. SUMMARY AND DISCUSSION

This work shows that current information and communication technologies can be applied to stroke rehabilitation. It is important that the engineering teams involve users at the start of the system design and then to get feedback at regular intervals. Once prototypes have been developed these should also be evaluated by naïve users, since expert users may have become familiar with operation of the system. This process requires input for a multidisciplinary team that includes therapists, ergonomists and researchers familiar with qualitative research. We have also benefited from support provided by industrial partners and input from a Professional Design Team. The research demonstrates that it is feasible to apply the emerging motion sensor technology and information and communication technology (ICT), to develop a low-cost home-based system that could be used to support post stroke rehabilitation.

Acknowledgements: We would like to thank all the patients and carers who have given time to the project. We would also like to thank the support from Design Futures Sheffield, Xsens, Charnwood Dynamics Ltd. and The Stroke Association.

8. REFERENCES

DoH (2000), *National Service Framework for older people*. London.

GA Mountain, PM Ware, J Hammerton, SJ Mawson, H Zheng, R Davies, ND Black, H Zhou, H Hu, N Harris and C Eccleston, The SMART Project: A user led approach to developing and testing technological applications for domiciliary stroke rehabilitation, *3rd Cambridge Workshop on Universal Access and Assistive technology (CWUAAT 2006).*

HRW Miltner, H Bauder, et al (1999), Effects of constraint-induced movement therapy on patients with chronic motor deficits after stroke, *Stroke*, **30**, pp. 586-592.

PM Rossini, C Calautti, et al(2003), Post-stroke plastic reorganisation in the adult brain, *The Lancet Neurology* **3**, pp. 493-502.

H Zheng, ND Black and ND Harris(2005a), Position-sensing technologies for movement analysis in stroke rehabilitation, *Medical & Biological Engineering & Computing*, **43**, *4*, pp. 413-420.

H Zhou and H Hu (2005), Inertial motion tracking of human arm movements in home-based rehabilitation, *Proceedings of IEEE Int. Conf. on Mechatronics and Automation,* Ontario, Canada, 29 July – 1 August, 2005, pp. 1306-1311.

H Zheng, RJ Davies and ND Black (2005b), Web-based Monitoring System for Home-based Rehabilitation with Stroke Patients, *Proc. of 18th IEEE International Symposium on Computer-Based Medical Systems* (CBMS 2005), Ireland, pp. 419-424.

ICDVRAT 2006

Session VII. Visual Impairment

Chair: Ali Al-khalifah

Non-visual feedback for pen-based interaction with digital graphs

S A Wall[1] and S A Brewster[2]

Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow
17 Lilybank Gardens, Glasgow, G12 8QQ, UK.

[1]*steven@dcs.gla.ac.uk* [2]*stephen@dcs.gla.ac.uk*

www.tactons.org

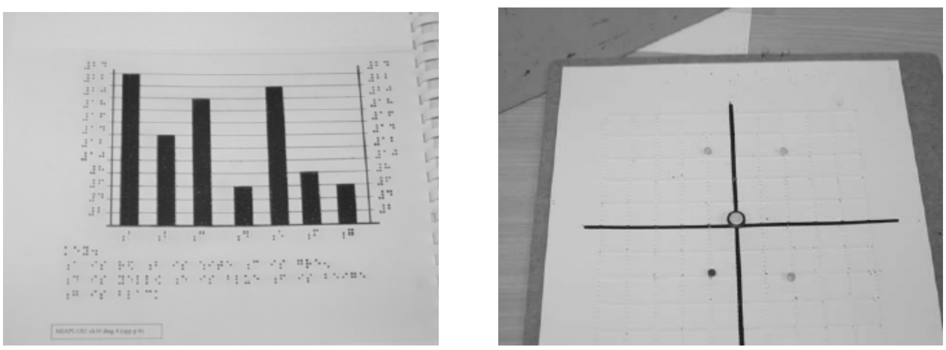
ABSTRACT

Access to simple visualisations such as bar charts, line graphs and pie charts is currently very limited for the visually impaired and blind community. Tangible representations such as heat-raised paper, and inserting pins in a cork-board are common methods of allowing visually impaired pupils to browse and construct visualisations at school, but these representations can become impractical for access to complex, dynamic data, and often require a sighted person's assistance to format the representation, leading to a lack of privacy and independence. A system is described that employs tactile feedback using an actuated pin-array, which provides continuous tactile feedback to allow a visually impaired person to explore bar charts using a graphics tablet and stylus. A study was conducted to investigate the relative contribution of multimodal feedback (tactile, speech, non-speech audio) during typical graph browsing tasks. Qualitative feedback showed that the participants found it difficult to attend to multiple sources of information and often neglected the tactile feedback, while the speech feedback was the most popular, and could be employed as a continuous feedback mechanism to support graph browsing.

1. INTRODUCTION

Access to data visualisations such as line graphs, bar charts and tabular data is highly problematic for many visually impaired people. People with low vision may find it difficult to apprehend data that is presented as a graph in a webpage, spreadsheet or other document. Lack of access imposed by current teaching materials can be a barrier to visually impaired people who wish to pursue studies in numerate disciplines, such as maths, the sciences and economics, and also these sectors are deprived of potential students and employees from the visually impaired community.

The most common techniques that are employed to allow non-visual browsing and construction of graphs in schools are raised paper diagrams and pin-boards. Heat-raised paper (Figure 1a) can be employed to generate a tangible representation of monochrome graphics, by printing a suitably formatted representation, and passing it through a special heater. Pins can be stuck in a cork board to represent data points, and joined by rubber bands in order to represent graphs and charts (Figure 1b). These "low-tech" representations are extensively used within the education of blind and visually impaired students, but suffer from several potential drawbacks.



(a) *(b)*

Figure 1. *(a) Raised paper bar chart, (b) pins in cork-board bar chart.*

Cues that are discriminable by vision may not always be easily discriminated by touch, and care must be taken not to excessively clutter tactile diagrams with information (for example, Braille legends and labels cannot be made “small print”, like visual text, due to the limits of tactile acuity). For these reasons, the assistance of a sighted person with specialist knowledge of how to prepare and format the tactile representation is required, leading to a lack of independence and privacy for the visually impaired person. Low-tech tangible representations of data are inherently non-dynamic and of a low resolution: this can make certain data representations such as complex line graphs or pie charts difficult to produce using a pin-board approach. There is also a lack of access to digitally stored data, which is particularly salient given the vast amounts of information available over the internet, which is largely inaccessible to the visually impaired community. Tangible representations are also difficult to store and subject to wear-and-tear; a student’s work will often have to be destroyed at the conclusion of a class, so that the materials can be reused by another student, which can also be very discouraging for a student wishing to achieve in a numerate discipline. Beyond the classroom, approaches to graph construction such as cork-boards and pins become entirely insufficient for day-to-day use in most professions due to their cumbersome use and lack of resolution.

Screen readers are the most common accessibility aid used by visually impaired people to access information with a computer. This software converts text from the desktop (e.g. icon labels, menu items, message boxes) or from a document, such as a web page, word processor document, or spreadsheet in to a stream of synthetic speech. Tangible representations of visualisations possess several advantages over a screen reader based representation. They support rapid, non-sequential browsing of data using the whole of both hands (although this is indirect, through inference based on relative heights of bars, paths of lines, or other features, dependant on the particular representation), combined with a pictorial representation that promotes a representation of the data analogous to that employed by sighted colleagues. When seeking to create a computer-based tool to support browsing of graphs, it would be beneficial to obtain the benefits of a digital tool, while preserving the advantages of representations in tangible media. One means of achieving this may be to use virtual reality technologies such as haptic interfaces or tactile displays that allow for development of dynamic representations of data that are accessible to visually impaired people using the sense of touch.

The “Tactons” project (*www.tactons.org*) is researching the application of tactile pin-array technology in the area of accessibility of graphs for visually impaired people. Tactile pin-arrays present dynamic information to the sense of touch using distributed mechanical deformation of the skin, usually over a small area such as the fingertips. Wall and Brewster (2006) previously described a system that employs tactile pin-array feedback for pen-based interaction with visualisations. Multimodal (haptic and audio) representations have been shown to perform better than force feedback technologies alone (Yu and Brewster, 2002). This paper investigates the use of speech feedback and sonification in conjunction with tactile pin-arrays. The aim of this study was to suggest guidelines for refining the interface, and for developing multimodal accessible solutions in general, and also to suggest areas where more empirical work is required in order to investigate combinations of tactile and audio feedback.

2. PREVIOUS RESEARCH

Work on using the sense of touch as an aid to computer accessibility has largely been focussed in three areas: augmentation of low-tech representations, force feedback technologies and tactile displays.

2.1 Augmenting low-tech representations

Tangible diagrams have been enhanced through the use of digital technology, for example, the “Nomad” system (Parkes, 1988) used a touch tablet in conjunction with a raised paper diagram of a map. The tactile information was supplemented with audio and speech cues that could be triggered by pressing on the tablet. More recently, the T3 system has applied this technique to create teaching media, including a world atlas, where a student interacts with a tactile raised paper map in order to obtain information about countries of the world via synthetic speech (Wells and Landau, 2003). This approach combines the benefits of being able to obtain a pictorial overview of data through the sense of touch, with the ability to discern detailed information provided by synthetic speech. A drawback of these approaches is that to create new content, the tangible raised paper diagrams still need to be formatted and produced, which often requires a sighted person’s assistance.

2.2 Force feedback

The first example of the use of force feedback devices to present visualisations to the blind was reported by Fritz and Barner (1996). Results from the Multivis project (*www.multivis.org*) showed that blind people were able to apprehend visualisations such as bar charts and line graphs presented using force-feedback devices, and answer questions regarding the data presented (Yu and Brewster, 2002). One of the main drawbacks of force feedback devices is that the user is denied the rich, spatially varying cues that are obtained when exploring a tactile diagram with the whole of both hands. Perception of shape is slower and more memory intensive, as the user must integrate temporally varying cues in order to build up an impression of the scene (Jansson, 2000). Further, many of the more sophisticated force feedback devices are very expensive for an individual.

2.3 Tactile Displays

Tactile displays present information to the user's skin via one or more smaller actuators. The VTPlayer mouse (*www.virtouch2.com*) is a commercially available, mouse based device that incorporates two tactile arrays, each consisting of a 4 by 4 matrix of individually controllable pins that deliver stimulation to the fingertips (Figure 2). The pins can be raised or lowered, but do not provide any resolution between this. During standard operation, the state of the pins is controlled by the pixels directly surrounding the mouse pointer. Using a simple threshold, a dark pixel corresponds to a raised pin, and a light pixel corresponds to a lowered pin. The user rests their index and middle fingers on the arrays while using the device like a standard mouse, and can feel a tactile representation of images. In this fashion, a blind user could potentially interpret the tactile cues and use them to navigate about a desktop environment, document or user interface.

Figure 2. *The VTPlayer tactile mouse (www.virtouch2.com). The mouse incorporates two, 4-by-4 tactile pin-arrays, on which the index and middle fingers are placed during operation.*

Jansson and Pedersen (2005) studied the performance of visually impaired users browsing a map with the VTPlayer. They observed that tactile information which indicated the crossing of borders on the map had no effect on performance in a navigation task, when used to supplement audio cues. The visually impaired users had many problems using the mouse. Mouse use is very difficult without any contextual information on target location when moving. The tactile arrays are too small to allow movement planning; further, in the absence of continuous visual feedback, the effects of unintentional rotations or lifting of the mouse go un-noticed. Wall and Brewster (2006) also observed in studies with the VTPlayer that the size and resolution of the array was too small to allow the users to plan their movements in a manner analogous to that of a sighted user using visual information. Jansson et al. (2006) conducted a second map-browsing study with the VTPlayer, which investigated the effect of providing textures (patterns of raised and lowered dots) within the borders on the map. They found that provision of texture had a significant effect, which was detrimental on the time to attain targets, surmising that the textures interfered with detecting crossing of borders.

3. SYSTEM DESIGN AND DESCRIPTION

3.1 System Overview

Following requirements capture conducted with visually impaired students at the Royal National College in Hereford, UK, a prototype system for exploring visualisations with a tactile pin-array was developed and evaluated (Wall and Brewster, 2006). The prototype system (Figure 3) uses a Wacom Intuos-2 graphics tablet for user input. The user provides input to the system using the tablet's stylus with their dominant hand, by moving the stylus on the active surface of the tablet, which controls a pointer in the application software. A VTPlayer mouse is used to provide continuous tactile feedback using the device's pin-arrays, but the mouse input is disabled, hence the user rests their non-dominant hand passively on the mouse, with the index and middle finger on the VTPlayer's tactile displays. The VTPlayer's pin-arrays provide tactile feedback to inform the user whether or nor they have the pen positioned over a bar. Audio output is provided by synthetic speech feedback and sonification.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

Figure 3. *The user browses the graph using the graphics tablet for input with the dominant hand, and the VTPlayer tactile mouse for output with the non-dominant hand.*

3.2 Graphics Tablet

Previous research had identified that the majority of visually impaired users were uncomfortable with using a mouse as a pointing device. Therefore, input is provided by a Wacom graphics tablet with stylus (*www.wacom.com*). The stylus is held with the dominant hand and used to control a pointer in the application, while the non-dominant hand is rested passively on the displays of the VTPlayer. The graphics tablet operates as an absolute positioning device, which allows the user to more effectively employ proprioception as a means of continuous feedback, compared to when using a mouse. By moving the stylus to point at the same physical location on the tablet, the application will receive the same co-ordinates as an input. This is not the case with a mouse, which can be unintentionally lifted or rotated, creating an inconsistent mapping between the physical and virtual spaces.

When exploring a simple bar chart, the X and Y axes are a constant point of reference for the user, therefore the tablet was augmented with tangible X and Y axes (see Figure 3), providing an unambiguous position reference which the user could quickly identify to guide and direct their subsequent exploration of the graph. Creating a physical representation of the graph's axes allows the user to quickly locate and disambiguate them from the graph's data.

3.3 Tactile Feedback

The software application takes a file containing a list of values and corresponding labels as input, and renders a graphical representation of a bar chart (Figure 4a). The VTPlayer renders tactile information based on the corresponding pixel information surrounding the mouse pointer, as the user explores the graph by moving the stylus. The pins of the device are raised as the stylus moves over a bar and lowered when the stylus is not over a bar. The bars are not patterned so as not to unnecessarily clutter the tactile representation with shapes and textures, which are confusing for users (Wall and Brewster, 2006, Jansson et al., 2006). The status of the VTPlayer pins can therefore be employed for navigational purposes ("Am I on a bar?") or to estimate values of bars indirectly ("How high is the bar?").

Figure 4. *(a) Representation of a graph within the software, and (b) raised paper version.*

3.4 Speech Feedback

The stylus can also be used to instantiate speech feedback by tapping the tip on the surface of the tablet. This triggers some contextual speech feedback that tells the user whether they are (a) outside the graph, and if so, in which direction, (b) on a bar, and the label and value of the bar, (c) above a bar, and the name of the bar, or (d) in-between two bars, in which case the name of both bars are given. Preferences for parameters of the speech, such as speed, or the voice that is used, can be set by the user through the windows control panel.

3.5 Sonification

By clicking the stylus button, users can hear a sonification of the data. For each bar in the graph, in order, from left to right, a MIDI piano note is played, with the pitch proportional to the data value. The purpose of this was to allow the user to obtain a quick, auditory overview of the graph (e.g. where particularly high or low areas were), such that this could be used to subsequently direct their querying of the data values. The notes were played sequentially at 200 millisecond intervals using the general MIDI acoustic grand piano instrument (instrument 0). The highest value in the graph corresponded to MIDI note 100, the lowest value in the graph corresponded to MIDI note 35, as recommended by Brown et al. (2003), with the rest of the notes scaled linearly in-between.

4. EVALUATION

In order to evaluate the visualisation system, a usability study was conducted with visually impaired computer users at RNC Hereford, in November 2005. The purpose of the evaluation was to gather feedback from the users in order to identify guidelines to further refine the design. As the system was at a prototype stage, evaluations were qualitative, using a think-aloud methodology and post-hoc interviews with participants. Specifically, we were interested in finding out about how the participants would use the different output modalities of speech, tactile and non-speech audio feedback to help them navigate and comprehend the visualisation, using a series of probe questions.

4.1 Participants

Ten visually impaired participants took part in the evaluation, nine were students at RNC and one was a member of staff at the college. All were studying, or had studied to a level of “further education” (e.g. A-levels, B-Tec., etc.) Some of the subjects were congenitally blind, while others were adventitiously blind. The subjects had various levels of residual vision. Due to the limited subject pool we decided not to control for these factors in the evaluations. The evaluation lasted approximately one hour per participant.

4.2 Experimental Procedure

The first stage in the experiment was to introduce participants to the graphics tablet and the VTPlayer tactile pin-array; during this stage the experimenter explained how to use the devices and how to interact with the system. The participants then undertook a training exercise intended to give them some insight in to how the tablet and tactile display system worked together. This training introduced the participant to the representation of graphs, using both the tactile pin-array and audio feedback. This was introduced by way of analogy to a raised paper diagram. The participant was handed a raised paper version of a bar chart consisting of twelve bars. This representation of the graph did not have any labels or legends, and was exactly as the representation appeared in the application software (Figure 4b).

The experiment itself used example graphs with verbally delivered probe questions from the experimenter in order to encourage the participant to use the tools and techniques available to them, and to stimulate their opinion on the relative contribution of the different interface components. Analysis was based on user comments, as the interface was still in a prototype stage. Four graphs were shown to the participants, and the corresponding questions (and answers) were as follows. Graphs were based on meteorological data (*www.worldclimate.com*):

1. Average temperature per month (in degrees centigrade) in Bangkok: Describe the overall shape of the graph (rises to a peak in July and then decreases), what are the hottest and coolest months (July and January)?
2. Average rainfall (mm) for various cities of the world in April (12 cities): Which has the most rain out of Brisbane, Geneva and Milan? (Brisbane) Is it wetter in Christchurch or San Francisco? (Christchurch) Which city has the highest rainfall overall? (Auckland) If there was twice as much rain in Copenhagen, would it be wetter than Geneva? (Yes).
3. Average rainfall (mm) per month in Brisbane: Describe the general trend of the graph (Starts high in January, decreases to a minimum in August and September before increasing again up to December). What were the highest and lowest months for rainfall? (January and September, respectively).
4. Average rainfall (mm) for various cities of the world in October (12 cities): Which has the most rain out of Brisbane, Copenhagen and Paris? (Brisbane) Is it wetter in Amsterdam or Venice?

(Amsterdam) Which city has the highest rainfall? (Milan) If there was twice as much rain in Budapest would it be higher than Brisbane? (No)

Finally, the participant took part in a post-hoc interview with the experimenter. With the participant's consent, the entire training and evaluation process was recorded using a portable MP3 player with built in microphone. The experimenter also made notes during the evaluation of user behaviours and strategies that were observed.

5. RESULTS AND DISCUSSION.

The recordings of each participant were transcribed and analysed in order to derive common preferences and recommendations for improving the usability of the interface. The key findings were as follows:

5.1 Speech feedback was the most preferred modality of the participants.

All the participants relied extensively on the use of speech feedback in order to answer the questions. Three participants openly remarked that this was because they were very familiar with speech representations due to prior experience with screen readers. Two of these three participants also commented positively on the fact that the speech gave them an exact value to several decimal places, which would be difficult to estimate when using a tactile diagram on raised paper.

- *"I did that because I knew the speech was there … that's just what I've been doing for so long", P1*
- *"… you can obtain that information a lot faster with the speech than you could before (without speech) because it was sometimes difficult to feel it when there was only a slight difference", P2*
- *"… you've got the height there with the speech, you just know it, you're just obtaining the amount and the highest bar absolutely in one, immediately.", P2*
- *"…having the facility of speech… I'm taking advantage of that and relying on it to its full potential" – P6*

As requesting the speech information was such a common action, the means of obtaining the information (tapping the pen) had to be efficient and free of errors. Five of the ten participants wanted the option to receive this information automatically while browsing the graph, without having to tap the pen or click a button. Lifting the pen to tap occasionally caused errors in positioning when placing the pen back down. Similarly, clicking the stylus button occasionally caused the pen to slip on the surface of the tablet.

- *"I still think for someone like me it'd be better to get speech straight away rather than tap the pen or press a button", P9*

5.2 More control is needed over the sonification.

In comparison to the speech feedback, the participants did not often use the sonification. Five participants highlighted a lack of control over the sonification and that it played too fast for them to be able to apprehend the information. This made comparison between values difficult. It was often the case that a participant would want to compare a subset of the values, but would have no option to do so, without having to listen to the entire sonification again.

- *"… and then you could think about it a little bit, but it's (imitates sonification in exaggerated manner), and it takes about 2 seconds … where as with the speech you can do it in your own time. So you're losing the control over how fast you want the information.", P2*
- *"It would be better if you could control the speed, this would give you an overall view at that speed, but if you slowed it down you could pick out the specific months", P7*
- *"I think it was a little bit too fast because I got a bit confused. I was able to get a good picture of how they changed, but I think I would be able to get a better picture of where the changes were if it was a bit slower.", P8*

Four participants specifically stated that they thought playing the notes in response to their movements with the pen (e.g. when they move over a bar, play the corresponding note) would allow them to control the speed of the presentation, and also to select a subset of the values to explore more easily.

- *"it's such a quick sound, it's not sort of accurate, but if I could go over it and only make one note per bar it would be easier that way", P6*
- *" … if you could move your pen like this and you could move across and go (imitates sonification) … to get more control over the speed of the notes, and if you weren't sure of the difference between the*

two then you could quickly check by putting the pen on those two. You choosing the notes would make it sink in better because you're actually making the sound happen", P8

5.3 *Tactile feedback was often neglected by the users.*

Four of the participants often forgot to place their non-dominant hand on the tactile displays of the VTPlayer. When probed by the experimenter about this, it was most often the case that the participants found it difficult to attend to multiple sources of information.

- *"When you're using the mouse and the speech and the pen it's a lot to take in at once", P2*
- *Experimenter: "I notice you haven't got your hand on the tactile?"*
- *P3: "I was too busy listening".*
- *"I was so focussed on here I wasn't paying attention to the mouse", P7*

The speech feedback gave enough contextual information to allow the participants to navigate without using the tactile feedback. Participants were able to explore the graph and apprehend the information needed to answer the questions by relying on the speech feedback to control the pen movement. Essentially, the speech was used as a means of continuous feedback to control the pointer.

- *"… the good point is that the speech was guiding me all the time and giving me several lots of information", P1*

One participant suggested that mounting the tactile feedback in the pen might free them from having to attend to too many separate sources of information. This would also free the non-dominant hand and allow it to be used for place-marking, guidance, or to gain more stability with the pen.

- *"I was thinking, well you're holding the pen, the nib, if you're holding the nib when you reach the top of the graph, all of a sudden the nib comes up, 'right that's the top' … well, yes, you'd feel it in some way on your finger … if you could incorporate it in to the pen you wouldn't need two separate devices.", P7*

5.4 *Obtaining multiple data labels using speech was error prone and time consuming.*

When working with nominal data on the X-axis (for example, the cities of the world in questions 2 and 4) the participants would often explore all the bars in a graph before attempting to answer a question. This could be quite a time consuming task and prone to errors, due to the number of movements of the stylus and taps that were required to access the speech information. The experimenter noted that it was confusing for the participant if they thought they had traversed to the next bar, but instead tapped on the same bar again, or the pen slipped and instead tapped between two bars.

5.5 *Obtaining an overview of the data through tactile feedback was time consuming.*

Three of the participants commented that access to data through the tactile feedback was currently impossible given the absence of any scale, gridlines or markers that could be used to help estimate the values. Suggestions included displaying gridlines or scale through tactile information, or through audio, using subtle sound cues.

- *"… every ten you had a dot, it was going "beep". If you was going up the graph from the X-axis and every sort of ten millimetres there was a piano key, you'd know where you were, wouldn't you?" – P7*

6. CONCLUSIONS AND FUTURE WORK

In summary, the multimodal output provided sufficient feedback to allow the users to interact with the system. A number of possible refinements were derived through analysis and consideration of the results of the usability evaluation.

The speech feedback was extensively used by all the participants. For all of the participants, and many blind computer users, screen readers are a part of their daily life, and a familiar technology. Speech feedback can rapidly deliver exact information without the need to encode/decode in another modality. As speech is very commonly used, access needs to be quick and error free. User feedback indicates that tapping with the stylus, or clicking the stylus button might not be the most appropriate means of requesting speech information, therefore these interactions could be investigated to quantify error rates and accuracy. Speech feedback can be used to present resources such as legends, axis titles and labels for nominal data series, but these need to be implemented as "one click" actions, in order to reduce the possible errors. One possible solution would be for the user to click or gesture on the relevant axis to receive the information.

The tactile pin-array feedback was neglected by the users. One solution to attempt to alleviate this would be to avoid multiple sources of stimulation and control. For example, tactile feedback could be delivered through the stylus, or by vibrating the tablet, rather than by having the user attend to a separate information source with their non-dominant hand. Tactile feedback could be used to provide subtle cues to aid with target acquisition and error retrieval (e.g. detecting slip-off target errors). In future work we will investigate and quantify the role of tactile feedback in non-visual target acquisition tasks. Tactile feedback can also be better supported and given context by employing tangible media, such as the tangible X and Y axes in the application as described. Tactile display provides dynamic feedback, but tangible media can be used to give context and meaning through "the bigger picture", as it can be rapidly explored with the entire hand. As an example, tangible grid lines could be employed on the graphs to communicate scale, provide navigation aids, and to allow the user to use the dynamic tactile feedback to estimate bar heights.

Allowing more personalization and control may make the sonification more usable. Being able to control the speed of delivery and being able to filter the information to make comparisons more easily were the most common requests. Supplementing proprioceptive and kinesthetic feedback with audio feedback, by playing the notes in response to the users movement may support apprehension of data through this technique. This will be implemented and tested in a future version of the system.

These guidelines will be employed to develop further iterations of the prototype system, and may be useful to designers developing multimodal solutions for other accessible applications, such as web browsing or gaming.

Acknowledgements: We wish to thank RNC Hereford for access to students for participation in the evaluations. This research was supported with funding from EPSRC grant GR/S53251/01.

7. REFERENCES

L M Brown, S A Brewster, R Ramloll, M Burton and B Riedel (2003), Design guidelines for audio presentation of graphs and tables, *Proc. ICAD 2003 Workshop on Auditory Displays in Assistive Technologies,* Boston, pp. 284-287.

J P Fritz and K Barner (1996), Design of a haptic graphing system, *Proc. 19th RESNA Conference*, Salt Lake City.

G Jansson (2000), Basic issues concerning visually impaired people's use of Haptic displays, *Proc. 3rd International Conference on Disability, Virtual Reality and Associated Technologies,* Alghero, Sardinia, Italy, pp. 33-38.

G Jansson and P Pedersen (2005), Obtaining geographical information from a virtual map with a haptic mouse, *Proc. XXII International Cartographic Conference (ICC2005),* A Coruna, Spain, Available on conference CD-ROM.

G Jansson, I Juhasz and A Cammilton (2006), Reading virtual maps with a haptic mouse: effects of some modifications of the tactile and audio-tactile information, *The British Journal of Visual Impairment*, **24**, *2*, pp. 60-66.

D Parkes (1988), "Nomad": An audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people, *Proc. of the Second International Symposium on Maps and Graphics for Visually Impaired People,* London, pp. 24-29.

S Wall and S A Brewster (2006), Feeling what you hear: tactile feedback for navigation of audio graphs, *Proc. of ACM CHI 2006,* Montreal, Canada, pp 1123-1132.

L R Wells and S Landau (2003), Merging of tactile sensory input and audio data by means of the talking tactile tablet, *Proc. Eurohaptics 2003,* Dublin, Ireland, pp.414-418.

W Yu and S A Brewster (2002), Comparing Two Haptic Interfaces for Multimodal Graph Rendering, *Proc. of IEEE VR2002, 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (Haptics 2002),* Florida, USA.

Developing an ENABLED adaptive architecture to enhance internet accessibility for visually impaired people

C C Tan[1], W Yu[2] and G McAllister[3]

[1,2]Virtual Engineering Centre, [3]Sonic Arts Research Centre,
Queen's University Belfast, Malone Road, Belfast, UK

[1]*s.tan@qub.ac.uk*, [2]*w.yu@qub.ac.uk*, [3]*g.mcallister@qub.ac.uk*

[1,2]*www.vec.qub.ac.uk*, [3]*www.sarc.qub.ac.uk*

ABSTRACT

This paper gives an overview of the current status of Internet accessibility and offers a brief review of the existing technologies that address accessibility problems faced by visually impaired people. It then describes an adaptive architecture which is able to integrate diverse assistive technologies so as to allow visually impaired people to access various types of graphical web content. This system is also capable of adapting to user's profile and preferences in order to provide the most adequate interface to the user.

1. INTRODUCTION

We rely very much on the Internet to acquire a wealth of information on different subjects and areas in our daily life. Due to the ability of the Internet in making many kinds of information and services easily accessible, it subsequently turns the Internet into a disputable communication and information medium that can be accessed by everyone. However, people with visual impairments have been prevented from taking the full benefit of this prominent new technology due to the accessibility issues and barriers attributed to the predominantly visual content of the Web (Hackett et al, 2004). One of the disadvantages faced by blind people is that, they are not able to fully access graphics available on the Internet, which might carry a comprehensive set of information. They can only get a general idea about the graphics from the text descriptions provided within the graphic via speech output systems, for example a screen reader, or Braille display, provided that the text descriptions are available and precisely describe the graphic. In order to improve graphics accessibility, much research has been performed in an attempt to present information embedded in graphical Web content to visually impaired people. Many studies showed that haptic feedback can serve as a substitute for visual feedback, due to the high similarity between visual and haptic modality (Keyson, 1996; Gibson, 1966). For instance, Picard et al (2003) has used both cross-modal matching and transfer task to show that vision and touch were perceptually equivalent for natural textures perception. Henceforth, numerous haptic technologies and research related to this area have been introduced and carried out respectively in accordance with the results of these studies. For instance, an online virtual graph construction tool which allows blind people to create and explore a virtual graph using a force feedback mouse has been developed by Yu et al (2003). TRIANGLE, a computer program developed by Gardner et al (1998) uses non-speech audio as an alternative technique to traditional graphing methods.

Despite that there is so much research aimed to boost the accessibility of graphical content on the Internet for visually impaired users, they provide their own solution to a particular problem based on a fixed set of hardware. For example, the applications only works with a force feedback mouse to explore graphs, a Braille display for text, or audio sonification methods to present computer games, etc. There is a lack of an integrated platform that allows visually impaired people to access any graphical Web content without worrying about the type of device to use. Therefore, it is important to have an interoperable system that is able to handle various existing assistive technologies and competent of allowing users with visual impairments to access a variety of graphical Web content.

This paper firstly describes the main objective of a European research project – ENABLED, and details an adaptive architecture that has been designed and developed as part of this project. This architecture is developed to be capable of incorporating different assistive technologies to access diverse types of graphic content on the Internet. As an adaptive system, this architecture also adapts to the user's need and enhances

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

user interaction with the system by considering user's profile and preferences. The paper is then followed by a short scenario that presents the use of the adaptive system. Finally, the paper concludes by presenting the potential impacts of this system architecture to people from various realms.

2. PROJECT DESCRIPTION

ENABLED (Enhanced Network Accessibility for the Blind and Visually Impaired) is an Integrated Project funded in the Framework 6 Programme and consists of 13 partners across Europe. The ultimate goal of the ENABLED project is to improve blind and visually impaired people's access to the information and services available on the Internet. Both desktop-based and mobile-based access are being investigated. In the mobile scenario, information about users' surrounding environment is provided on portable devices, for example a PDA, in order to improve their mobility. In order to achieve its objective, this project (1) develops technologies that create universal accessible content for the Web, and algorithms that convert existing inaccessible contents to be accessible; and (2) develops ubiquitous tools that enable easy access to information and interfaces that are adaptable and interoperable regardless of the location of the user and the equipment the user is using.

ENABLED is an ongoing project where one of its aims is to develop an adaptive interface for blind people to access different graphical content on the Web (Tan et al, 2005). The notion of developing such architecture is to create a system that is able to adapt according to various requirements and criteria. Therefore, it is developed to be interoperable and scalable by supporting multiple assistive technologies, is independent of graphical Web content and can adapt to a diverse range of user's profile and preferences. The rationale of this architecture is explained in further detail in one of the following sections of this paper.

2.1 Related Work

An adaptive system is a physical system that is capable of self-adapting its behaviour in response to its environment. The adaptation of a system is often relevant to achieving a goal or objective. It tries to adapt to relevant variables in order to achieve its goal. Normally, the goal can be realized in systems that have feedback (Bar-Yam, 2000). Adaptive systems have been considered in a wide range of recent research efforts, in which two of the most common domains are mobile computing (Sun and Sauvola, 2003) and hypermedia (Stephanidis et al, 1998a). The variables in which an adaptive system can adapt to vary according to the main objective it intends to accomplish, for instance users, environment, media devices, input/output devices, platforms, network services and so on. Nevertheless, adaptation to users has always been considered when developing an adaptive system, in accordance with the definition given by Oppermann (Oppermann, 1994), where an adaptive system should be able to change its own characteristics automatically according to the user's needs. Designing a user adaptive system alone can be one significant area of research. Individual user adaptation may include adapting their needs, goals, preferences, knowledge, interests, skills, impairments and so on.

Attempts have been made to achieve interoperability between different input devices. For example, Gammenos et al (2005) has developed UA-Chess, a universally accessible Internet-based chess game that uses a variety of alternative input/output modalities and techniques in any combination according to various user profiles. Meanwhile, another reference project, AVANTI, has developed and evaluated a distributed system that provides hypermedia information about a metropolitan area for a variety of users (Fink et al, 1997). In this project, the integration of input and output devices is provided by a Device Software Layer to uniformly control and communicate with different input and output devices (Stephanidis et al, 1998b). In order to adapt to existing Web content to meet the specific needs of user, Huang and Sundaresan (2000) has developed an extensible transcoding system, Aurora, to help the broadest population of users, particularly in the disabled community, to obtain various Web-based services, such as online auctions, search engines, etc. It uses a transaction model to characterize user's abstract goals, and then extracts and adapts only the Web content relevant to the goal in order to be presented to the user.

Basically ENABLED adaptive architecture possesses similar objectives to most of these systems, which is to create an adequate interface to its user by adapting to the user's requirements and favourites and other variables. However, in addition to adapting to user's needs and preferences, the ENABLED system also focuses on the adaptation of different assistive technologies that are available for people with vision difficulties so as to enable them to gain access to the common types of Web graphics.

2.2 System Architecture Overview

An adaptive system should be designed to be aware of and adaptable to different variables according to its aims. In the domain of mobile computing, research and studies are carried out to find models and techniques to design an adaptive application which is able to support different devices (for example cell-phones, PDAs), adapt to the environment (if it is noisy room, low light), or perform adaptation based on user requirements. The goal of this research is to develop an adaptive system which enables visually impaired users access to various graphical contents on the Web through existing assistive devices. Therefore, there are three main variables that are to be considered in this system, which include (1) various existing assistive technologies; (2) diverse types of graphical web content; and (3) user's profile and preferences. The fundamental concept of the ENABLED adaptive architecture is depicted in Fig. 1, which also illustrates the interaction between these variables and its end-users.

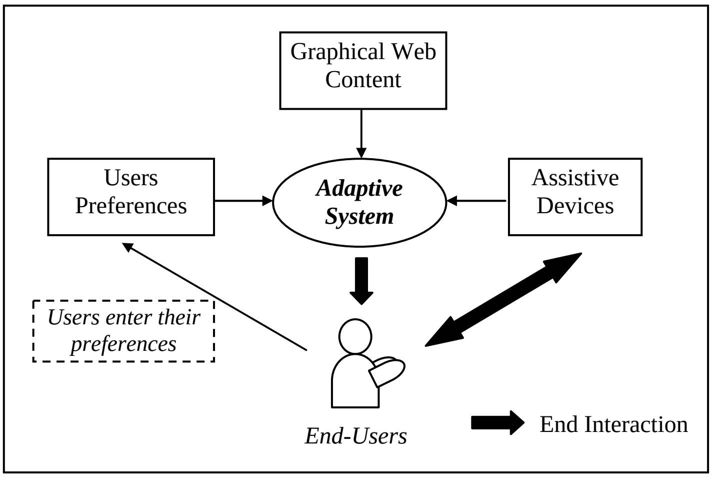


Figure 1. *Graphical presentation of the basic concept of ENABLED adaptive architecture, which adapts to assistive devices, graphical content and users.*

Nowadays, there is a wide range of assistive technologies available to improve Internet accessibility of users with visual difficulties: tactile or haptic modalities, 2D or 3D devices, low or high cost technologies and so on. Users are free to choose the assistive device they like or they could afford. Hence, it is vital to ensure the architecture is capable in adapting at a device level. On the other hand, as graphics have become one of the dominant components in Web pages, there is a wide variety of graphic information offered online. Examples of available Web graphics include information-oriented graphics such as graphs, interactive maps, images, drawings, photographs, and so on. These graphics can be in numerous file formats, for instance, the most commonly used computer graphics file formats on the Internet are JPEG, PNG, SVG and so forth. By considering the diversity of the graphics, users who utilize this architecture are not restricted to explore only certain type of Web graphics. An adaptive system often involves adaptation to its users. Interaction between user and system is crucial and the latter ought to be able to have the competency to meet the individual abilities, requirements, and favourites. In the following section, a detailed description of the ENABLED adaptive system that is capable of adapting to users, assistive devices and graphical web content will be presented.

3. SYSTEM ARCHITECTURE DESCRIPTION

This adaptive system is compatible to most of the Microsoft Windows-based browsers, including Windows Internet Explorer and Mozilla Firefox. This is one of the advantages of this system as statistics show that both of these browsers are the most popular and commonly used by Internet users (W3Schools, 2006). The architecture basically composed of five main components, namely: (1) Web Content System, (2) Context Manager, which consists of Configuration System and Preferences System; (3) Core Processor Module; (4) Application Database; and (5) End Interface. Each of these components is responsible for different roles and interacts closely with each other so as to constitute a seamless adaptive system. The interaction between all the aforementioned components and end-user are depicted in Fig. 2. The end-user provides input information about the graphical Web content and chooses the assistive technologies to be used for the graphic exploration. The information about the graphical content, which is the type and the file format of the graphic, are then perceived by the Web Content System. Adaptation to users is handled by the Preferences System, which is part of the Context Manager. Besides taking care of the user's requirements and favourites, Context Manager also presides over the computational and device technical information with the help of the Configuration System. By gathering the provided inputs, the Core Processor Module processes and administers this information and locates the most appropriate end interface for the user from the Application Database. The Application Database stores stand-alone sub-applications which are responsible for both graphics and haptic-audio rendering. The End Interface communicates with the assistive technologies, which could be an audio speech output and/or haptic modalities, so as to be presented to the end-users.

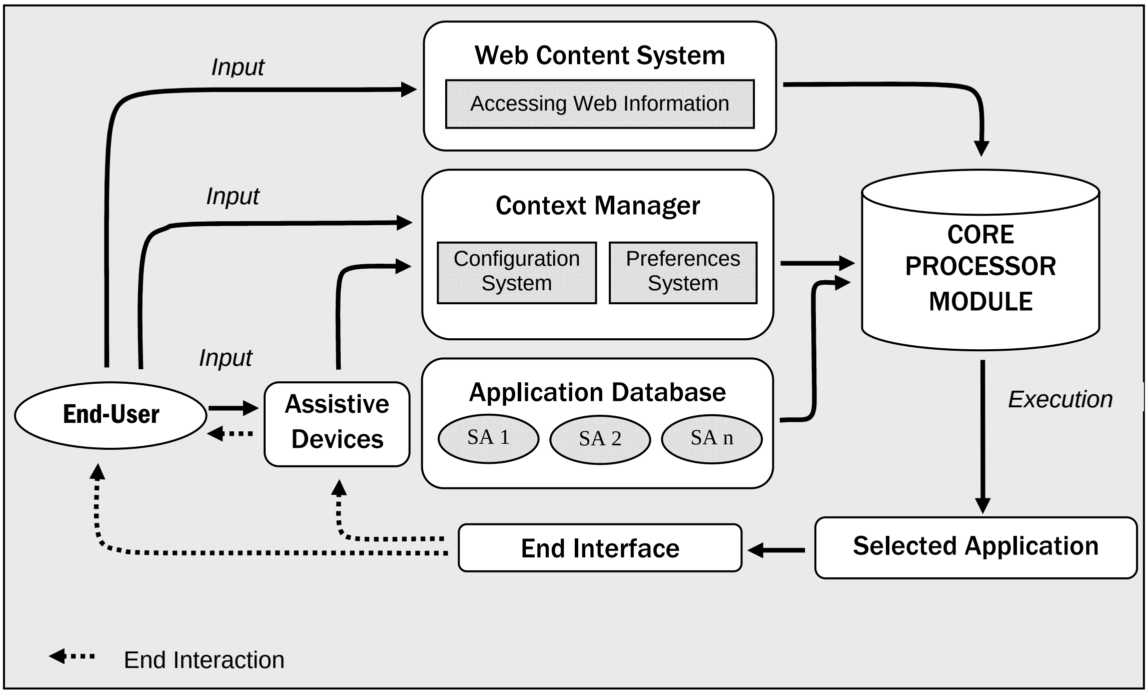


Figure 2. *Architecture of the ENABLED adaptive system.*

3.1 Web content detection and system launching

The first part of the system has been developed, which is to detect the types and formats of graphical Web content, and also to launch adequate interfaces for its users. The following sub-sections elaborate the underlying design principles of the developed system and interaction between the system and end-users.

3.1.1 Design considerations for the system launching. During the process of designing and developing, a number of significant design factors which are closely related to the end result of the use of the adaptive system have been considered. One of the apparent aspects that ought to be considered when designing a human-computer interaction system is the user interaction issue (Schmidt, 2000). The interaction between user and system should be simple, straightforward and understandable especially when the user target group for this adaptive system is people with vision impairments. The involvement of users should be kept to a minimum. Additionally, the main input and output devices that blind people use on a computer are the keyboard, Braille display or text-to-speech engines due to the use of mouse to navigate could cause frustration and disorientation to them (Yu et al, 2004). Therefore, the human-system interaction in this architecture can be accomplished by using keyboard or mouse independently. The restrictions faced by visually impaired users have been taken into account, for example, by reducing the number of pop-up or prompt boxes displayed to blind users. The problem of having prompt boxes is that, most versions of text-to-speech synthesizers are not capable of reading the content of these message boxes. They only read the title bar of the pop-up boxes, which in many cases, the title of the pop-up windows tell nothing about the message, and this may cause difficulties and confusion to people with visual difficulties.

3.1.2 Design rationale of the system launching. The ENABLED system is a stand-alone application. It is developed by adopting the features provided by Microsoft Active Accessibility (MSAA) technology, which was introduced by Microsoft to provide a standard programmatic access to user interface (UI) elements of Windows-based applications (Ford, 2004). By implementing this approach, the system is able to manipulate the UI elements that are currently accessed by users on a Web page. In order to have access to a graphic contained within a Web page, this adaptive system needs certain information about the graphic. Therefore collaborations from Web developers are required. Web developers who would like to allow visually impaired users access to graphical content embedded in their Web page need to provide an additional button, the *"ENABLED"* button for each of their graphics. These buttons carry information about the graphics types and formats, and also the data file of the graphics. Data file is a file that contains raw data and information about a graphic so as to create the graphic. It could be in many file formats. For example, a graphic with SVG extension contains a mixture of script and XML for defining the graphic. It is required by the sub-application in order to create a suitable interface to users. To differentiate a button from one graphic to another, it is essential for Web developers to include keyword(s) for each graphic on the button after the "*ENABLED*" text. For instance, if a button associates to a map of Europe, the text of the button could be "*ENABLED: Europe Map*".

3.1.3 Trigger action and its rationale. By installing this adaptive system onto their machine, users are able to run the system in the background whilst they browsing or surfing the Web. A user who is prepared to employ the adaptive system to explore Web graphics should understand that Web pages that integrate with this system would have buttons with "*ENABLED*" text on them. These buttons act as the *trigger switch* of the system.

As recommended by various multimedia accessibility experts (National Council on Disability Washington, 1998), in order to ensure that blind users have full access to the Internet, there should be a keyboard equivalent for each designed feature in any of the human-computer system, in which mouse access should never be the only method of access. Consequently, the launching of this system is developed with the reference of this accessibility rule, in which it can be in the complete trigger by keyboard for the benefit of blind users who are not familiar with the use of the mouse. For users who are familiar with the use of the mouse, once they decide to know more about a graphic, they could simply click on the button, and press the *ENTER key*. Alternatively, the key can be pressed whenever the keyboard focus is on the button. On the whole, this trigger action notifies the system that the user is now ready to know more about this graphic. For user who uses a screen reader to navigate the Web, when they are sure that the keyboard focus is on the button which refers to the graphic they are interested (via their assistive technologies), they should press the *ENTER* key once to confirm the selection of this graphic. If JAWS screen reader is used, to turn off the virtual cursor used by JAWS, user has to press the "-" on the num pad or "*INSERT + Z*" simultaneously. By pressing the *ENTER* key once again, it notifies the system to go for the subsequent stage.

3.1.4 Results of the trigger action. The notification from users leads to subsequent operations. The system firstly downloads the data file of the graphic from the server to a specified folder that automatically created during the installation of the ENABLED system. It also determines the type and format of the graphic and notifies the Core Processor Module about the findings. The information about the graphic is vital as to allow the adaptive system to make the correct choice in launching the suitable sub-application accordingly.

3.2 Sub-applications

Both functions of graphic rendering and the generation of the haptic-audio interaction are encapsulated in the stand-alone sub-applications. It could be developed independently by researchers and developers in computer graphic rendering or haptic interactive techniques, and subsequently integrated with this adaptive system. A sub-application can extract the essential information of a graphic from its data file and translate them into their internal representation; for example, the internal data for a graph would be its numerical data, title of the graph, name for both axes, and so on. It is important to create and support the geometric representation of the graphics such as lines, curves, surfaces, etc so as to be used to get the corresponding haptic and audio feedback by analysing the user cursor position regarding the geometric model of the graphic. The available sub-applications have to be registered with the adaptive system so as to be recognised and stored in the Application Database in the system. In order to do so, information about the new sub-application needs to be provided to the system and included into the database.

3.3 Adapting to the user's context

In order to be able for the adaptive system to launch the correct sub-application in accordance with the selected graphic, the Core Processor System carries out a number of specific tasks by accumulating and manipulating the crucial information which has been transmitted to it.

The Context Manager is a cache that stores all the context information about the users, operating systems, assistive technologies and the available sub-applications, which are taken care by the Preferences System and Configuration System. The information is grouped into four main categories: (1) computational information, which is about assistive technologies (haptic, tactile, sound, speech), web browsers, computer operating system (version, language) and so on; (2) user information, which explains about their visual and other impairments, haptic/tactile and audio experience, and it also helps to identify users with their unique identification; (3) user preferences, which includes their preferable language, interaction methods, input and output interface and also about their computer and Internet skills; and (4) application preferences, which gives the information about available sub-applications that are recognised by the adaptive system.

There are two stages of checking before the related sub-application is called (see Fig. 3). After the detection of the type and format of the graphic selected by the user, the system compares this information with the configuration system to look for the appropriate sub-application that is able to handle this graphic format. If the result denotes that there are two or more sub-applications stored in the database which is capable of handling the same type of graphic, the system then searches for the user's preferable interface from the Preferences System and compares it with the feedback offered by those sub-applications. There is a possibility that two sub-applications provide the same kind of feedback to their user. When this situation happens, the system informs the user about the features, functions and limitations of both of the sub-applications. The user could choose from the sub-application selection according to the interaction style and other utilities, for example haptic utility or zooming, panning, scaling and so on, that offered by the sub-applications.

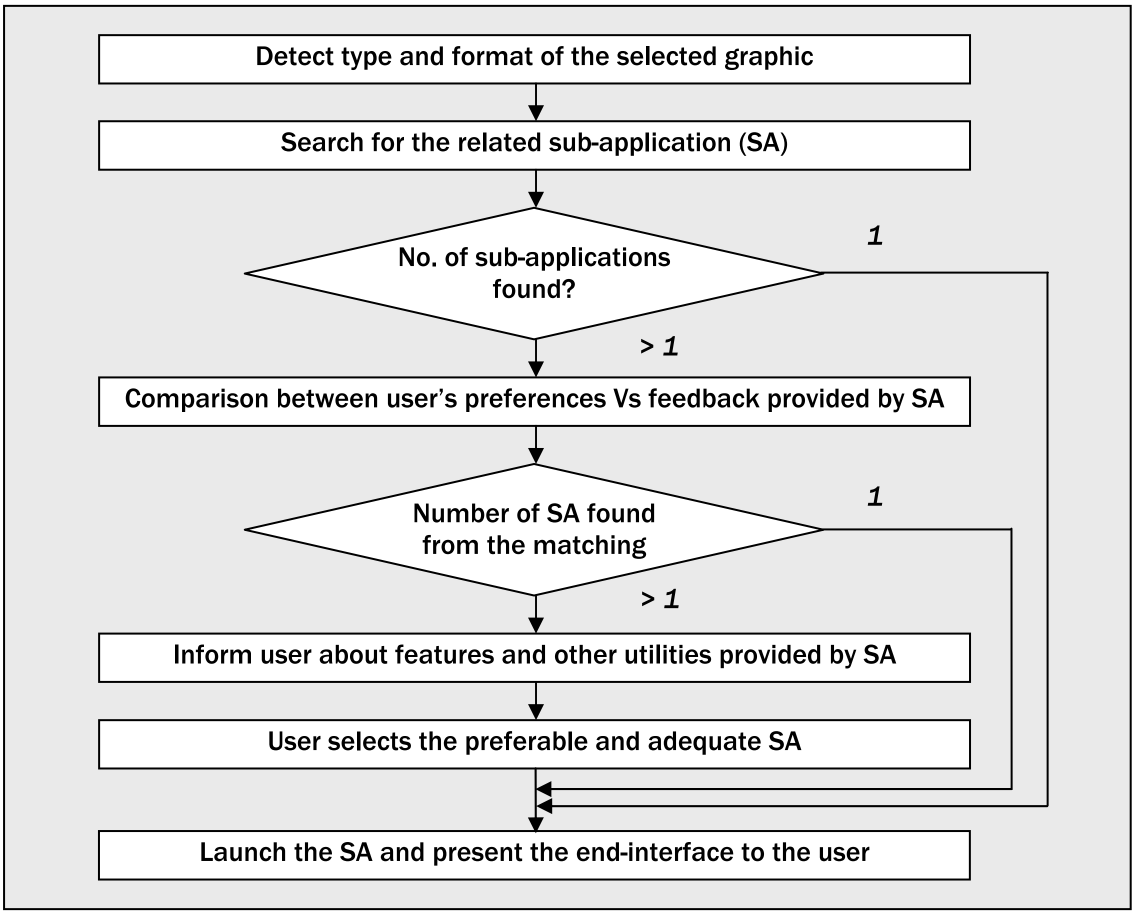


Figure 3. *Flow chart illustrates the process of launching the adequate sub-application which adapts to the graphic content and user's preferences.*

4. CONCEPTUAL EXAMPLE

A short scenario is depicted to present the use of the adaptive system described in this paper. Adam, who is a congenitally blind computer user, encounters a lot of graphs in his job. He decided to use the ENABLED system so that he is able to discuss with his colleagues about the graphs he found on the Internet. With the help of his colleague, the ENABLED system is installed on his computer and he is familiar with the specification of the system by reading through the readme file of the system, as well as from the online training on the ENABLED website (ENABLED online training website). He then creates his own profile and preferences with the system and sets them as the default settings.

Every morning, he runs the ENABLED system and also the screen reader application before he starts browsing the Internet using a standard Internet browser. As he is a JAWS screen reader user, Adam normally uses the up/down arrow keys and also JAWS quick navigation keys to navigate a Web page. When the keyboard focus is on a link or button, the text of these elements will be heard via his screen reader. One day, he found there are a few graphics on a Web page but he is only interested in one of them, which is a statistical graph indicating the trend in browser usage in 2006. By knowing that there is an ENABLED button for this graphic via his screen reader, he can differentiate that the Web page he is accessing is compatible with this adaptive system. Therefore, he presses the *ENTER* key once when he is sure that his keyboard focus is on the ENABLED button for that graph. Then he turns off the JAWS virtual cursor by pressing the num pad minus before he presses the *ENTER key again to start accessing the graph.*

The system then detects that it is a spreadsheet graph and downloads the excel data file of the graph to Adam's machine. From the Application Database, the system determines that there are two sub-applications which are able to handle an excel graph. On the other hand, from the Configuration and Preference system, the system identifies that there is no haptic device connected to the computer and the user prefers to have audio feedback and does not have any haptic/tactile experiences. Therefore, the system goes through the specification of both sub-applications and configures the features and feedback provided by them. As one of the sub-applications offers both haptic and audio feedback, whilst another provides only an audio interface, the system therefore launches the second sub-application and presents it to Adam. As a result, Adam is able to know more about the graph through his audio feedback and he is able to discuss this with his colleagues.

5. POTENTIAL IMPACTS OF ADAPTIVE ARCHITECTURE

The most noticeable merit that can be gained from the ENABLED adaptive architecture is dedicated to people with visual impairment, in which the system provides the ease of accessibility of Web graphical content to these users. Apart from the benefits it brings to the end-users, the system is able to incorporate different assistive technologies to increase the accessibility of graphics for users with visual difficulties. It is capable to make provision for the emergence of new assistive devices which might be introduced in the future. It gives the opportunity to adapt the system to future new developed functionalities and technologies. Software developers or researchers can develop new accessible applications based on this architecture in order to make more types of graphical contents accessible via different forms of assistive technologies. An assistive technology developer or a device manufacturer could easily integrate their interfaces with the ENABLED system by providing a hardware interface according to the ENABLED specifications. In additional, this architecture can act as a frame of reference to developers in other realms when they intend to develop a similar system of their own in their domains.

6. CONCLUSIONS & FUTURE WORK

In this paper, an adaptive architecture has been presented which is interoperable and capable in incorporating multiple assistive devices to gain access to diverse types of graphical Web content in different formats, and also adapting its user's context. The primary development of this system, which is to detect the graphical Web content that a user is interested in and to launch an appropriate sub-application, has been performed and presented in this paper. The fundamental concept and design of the continuing development of this adaptive system are also included as part of the paper.

This adaptive architecture has foreseen substantial benefits to people from diverse realms. Users with visual impairments would be able to gain the most advantage as the ENABLED adaptive system could improve their accessibility in exploring graphical web content and make such a task easier for them. With the existence of such architecture, it can save a lot of effort in developing systems that are capable of adapting to forthcoming devices and provide a significant solution to future needs. This adaptive system is believed to have the ability to enable more efficient interaction between the new assistive devices and the users, thereby increasing user experience. This is an ongoing project to show the feasibility of developing such system which is able to adapt and cope with several variables concurrently. Evaluation has been completed to assess the usability and accessibility of the developed prototype, which requires users to trigger an action to launch an adaptive interface. The results of the evaluation is being analyzed and further development on the architecture will be carried out in order to achieve the ultimate objective, which is to produce an adaptive and extensible system.

Acknowledgements. This work is supported by ENABLED – an Integrated Project funded by the EU Framework 6 Programme, IST (FP6-2003-IST-2-004778), www.enabledweb.org.

7. REFERENCES

A Schmidt (2000), Implicit human-computer interaction through context, *Personal Technologies*, **2**, 4, pp. 191-199.

C Stephanidis, A Paramythis, D Akoumianakis and M Sfyrakis (1998a), Self-adapting web-based systems: towards universal accessibility, *Proc. 4th ERCIM Workshop on "User Interfaces for All"*, Stockholm, Sweden, pp. 17

C Stephanidis, D Akoumianakis, M Sfyrakis, A Stergiou, N Maou, A Leventis,G Paparoulis and C Karagiannidis (1998b), Adaptable and adaptive user interfaces for disabled users in the AVANTI project, *Proc. of the 5th Intl. Conf. Intelligence in Services and Networks (IS&N '98)*, Belgium, pp. 153-166.

C C Tan, W Yu and G McAllister (2005), ENABLED adaptive architecture to improve Internet accessibility, *Proc. 2nd Intl. Conf. on Enactive Interfaces*, Genoa, Italy.

D Grammenos, A Savidis and C Stephanidis (2005), UA-Chess: A universally accessible board game, *Proc. Of the 3rd Intl. Conf. Universal Access in Human-Computer Interaction. G. Salvendy (ed.)*, Las Vegas.

D Picard, S Narcon, C Egoroff and L Kirsche (2003), Are vision and touch equivalent for natural textures perception? Studies from cross-modal transfer and matching tasks, *Proc. Of Eurohaptics 2003 Intl. Conf., Dublin, Ireland.*

D V Keyson (1996), Touch in user interface navigation, Doctoral Dissertation Eindhoven Technical University.

ENABLED online training website: http://www.training.enabledweb.org/

J A Gardner, R Lundquist and S Sahyun (1998), TRIANGLE: A Tri-modal access program for reading, writing and doing math. *Proc. of 1998 CSUN Intl. Conf. Technology and Persons with Disabilities*, Los Angeles.

J Fink, A Kobsa and A Nill (1997), Adaptable and adaptive information access for all users, including the disabled and the elderly, *Proc. of the 6th Intl. Conf. UM97*, Wien New York: Springer, pp 171-173.

J J Gibson (1966), *The Senses Considered as Perceptual Systems*, Houghton Mifflin, Boston.

J Z Sun and J Sauvola (2003), Towards a conceptual model for context-aware adaptive services, *Proc. 4th Intl. Conf. Parallel and Distributed Computing, Application and Technologies*, Chengdu, China, pp. 90.

National Council on Disability Washington (1998), Access to multimedia technology by people with sensory disabilities, Retrieved April 1, 2006, from http://www.ncd.gov/newsroom/publications/pdf/sensory.pdf

R Oppermann (1994), Adaptively supported adaptability, *Intl. J. Human Computer Studies*, **3**, 40, pp. 455-472.

S Ford (2004), Sara Ford's MSDN weblog, Retrieved February 20, 2006, from http://blogs.msdn.com/saraford/articles/223650.aspx.

S Hackett, B Parmanto and X Zeng (2004), Accessibility of Internet websites through time, *Proc. ACM SIGACCESS Conf. Accessibility and Computing*, Atlanta, Georgia, USA, pp. 32-39.

W3Schools (2006), Browser Information: Browser Statistic, Retrieved May 12, 2006, from http://www.w3schools.com/browsers/browsers_stats.asp.

W Huang and N Sundaresan (2000), Aurora: A conceptual model for Web-content adaptation to support the universal usability of Web-based services, *Proc. Conf. Universal Usability*, Virginia, US, pp. 124-131.

W Yu, G McAllister and G Dodds (2004), Multimodal Interfaces for Visually Impaired People to Access Internet, *Proc. 8th ERCIM Workshop – User Interfaces for All*, Vienna, Austria.

W Yu, K Kangas ands A Brewster (2003), Web-based haptic applications to allow blind people to create virtual graphs. *Proc. of 11th Haptic Symposium*, IEEE Press, Los Angeles, pp. 318-325.

Y Bar-Yam (2000), Concepts in Complex Systems: Adaptive, Retrieved April 28, 2006, from http://necsi.org/guide/concepts/adaptive.html

Exploratory strategies and procedures to obtain non-visual overviews using TableVis

J Kildal and S A Brewster

Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow
17 Lilybank Gardens, Glasgow, G12 8QQ, UK

johank@dcs.gla.ac.uk, stephen@dcs.gla.ac.uk

www.multivis.org

ABSTRACT

TableVis was developed to support computer users who are blind or visually impaired in tasks that involve obtaining quick overviews of tabular data sets. Previous work has covered the evaluation of this interface and its associated techniques of interactive data sonification and support exploratory processes. This paper examines the exploratory strategies and procedures employed by the users. A three-stage process for completing the exploratory task is described, and a discussion about the strategies and procedures that were observed is offered. Possible best practices and the most common issues are identified, which form the basis for the next steps to be taken in this line of research.

1. INTRODUCTION AND BACKGROUND

Advances in information technology in the last decades have contributed enormously to making information more accessible for users who are blind or visually impaired. With computers, accessible media can be generated dynamically in several forms and formats, which can often be produced without the need for assistance by the target users. Screen reading technology is an example of a particularly successful solution. The convergence of the advent of screen reading technology with the generalisation of personal computers for work, study and leisure, and the democratisation of information availability though the internet, has drawn a new scenario in which information is, by and large, far more accessible in non-visual forms. Other computer based technologies that can generate accessible content dynamically have also appeared, such as refreshable Braille displays, pin array-based tactile displays, and force feedback devices.

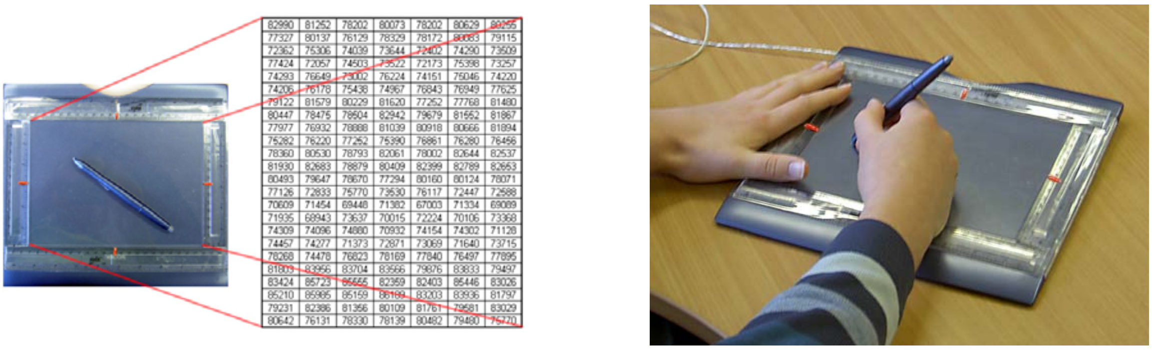
While these technologies provide access to computer based information non-visually, they do so presenting information in full detail, and there is little or no support to obtain overview information quickly and easily at the beginning of the exploration of a new data set. This can be a serious limitation when exploring information that is encountered for the first time. As Shneiderman defends (Shneiderman 1996), any data exploration starts by obtaining an overview, to be followed by zooming and filtering operations. To avoid overloading the user's working memory, detail information should only be obtained on demand, once an object of interest in the whole set of information has been identified. The accessibility technologies mentioned above provide fully detailed information directly, with no support to obtain initial overview information. Obtaining overview information non-visually is particularly difficult with numerical data sets, like spreadsheets. Accessing numerical values sequentially (as it happens when using screen readers) soon saturates the user's working memory (Miller 1956), and not enough comparisons can be performed between numerical values to reveal overview information. Not being able to obtain overview information from numerical data sets non-visually has serious implications for visually impaired students and professionals. They cannot freely choose to undertake higher education studies or career paths in disciplines that are numerically intensive, where collaboration regarding complex data sets is required, as well as keeping up with fast-paced data analysis.

Properties of human auditory perception have been successfully exploited to communicate data in sound (Kramer 1994). The concept of graphs in sound, obtained by mapping numerical values to pitch of sound was introduced by Mansur (Mansur 1985). Following this concept, substantial research work has been conducted in recent years in the field of auditory graphs, mainly with line graphs (Brown, Brewster et al. 2003; Walker and Cothran 2003). Extending these data sonification techniques, we have developed TableVis, a multimodal interface designed to browse tabular numerical data sets, with particular focus on providing means to obtain overview information quickly and easily, while still supporting filtering of information and obtaining details on demand at later stages in the exploration..

The aim of this paper is to get an insight into the exploratory strategies and procedures used to explore tabular numerical data sets using TableVis. Observed and recorded data from the explorations performed by blind and visually impaired users are analysed as a method of learning about the best ways of using the interface, common or unusual practises that should be further supported, and identify any common issues that might interfere with the exploratory process. Ultimately, this process of analysis could also provide some evidence about the mental processes of the participants in their strive to complete data exploration tasks.

2. PREVIOUS WORK WITH TABLEVIS

TableVis will is described in this section, mentioning only the functionality that is relevant to the discussion that follows. For a complete relation of functionalities available, requirements captured and a more detailed discussion about the iterative process that led to this design of the interface, the reader can refer to previously published work by the authors (Kildal and Brewster 2006a; b).



(a) *(b)*

Figure 1. *The data table to be explored is presented on the active area of the tablet, scaled to fill it completely (a).A user explores a data table creating interactive sonification with the pen, which the left hand feels the boundaries of the data set to provide contextual information (b).*

TableVis is a multimodal interface designed to explore tabular numerical data by generating sonifications of the data interactively, with particular focus on obtaining overview information. The interface was designed taking a user centred approach. Extensive requirements capture was conducted with users that were blind or visually impaired. Every iteration in the process of implementation and redesign was done combining qualitative and quantitative evaluations conducted with blind and visually impaired participants. In TableVis, two-dimensional data sets of any size or shape are scaled to fit on the whole active area of a WACOM graphics tablet (www.wacom.com), as seen in Figure 1, a. With the tablet's electronic pen, the user can point at any position on the tablet, which is equivalent to pointing at the cell that is mapped to that position on the tablet. This produces the whole row or column being pointed at to be sonified. Whether it is the row or the column that gets sonified will depend on the sonification mode that is selected at that time (these modes will be referred to as rows mode and columns mode). Sonifying a complete row or column involves sonifying each and every one of the cells in that row or column by mapping the values contained in them to pitch of sound (with higher values corresponding to higher pitches), and paying them from left to right (rows) or from top to bottom (columns), in very rapid succession. The succession is so fast that sounds are perceived to be simultaneous, as if they were played in a chord rather that in an arpeggio. The reason why sounds are played in an arpeggio rather than in a strictly simultaneous chord is that the user can modify the duration of the sound for each cell, stretching the sonification of a row or column in time, which functions as a detail filtering device as described in earlier publication by the authors (Kildal and Brewster 2006b); for the purpose of this discussion presented in this paper, however, it is sufficient to consider that all the sounds are perceived to be simultaneous. Presenting information in this fashion hides detail about the structure of the sounds in each row of column, and facilitates making simple comparisons between rows or columns regarding the average pitch of all the sounds contained in them. Thus, a user can tell which of two rows or columns contains sounds that on average have higher or lower pitch, or to identify outliers. Comparisons between all the rows or columns reveal features about the whole of the data set, such as trends and patterns along and across the data set. Exploring the same data table in rows mode and in columns mode presents two complementary representations of the same data object (these representations will be referred to as views of the data set, following a visual analogy). Each one of these views contains the same information (the same cells have rendered the same sounds), but grouped in two different ways that complement each other, like looking at the same object from two different angles. The interface does not present any information on the visual screen.

Several aspects of the design of TableVis are crucial for the users to be able to conduct non-visual explorations successfully. An absolute positioning pointing device (like a Wacom tablet in the case of TableVis) needs to be used for the user to be able to know about the position of the pointer in the data set. Additionally, the following invariants must be maintained to be able to keep both focus and context information throughout the exploration:

a. *The complete data set is on display.* The user knows that the complete data set completely fills the active area of the tablet, and that the physical, tangible borders of this active area also correspond to the boundaries of the data set.
b. *Information remains stationary in space.* The user knows that each cell in the table always occupies the same physical location on the tablet.
c. *The active area on the tablet has a fixed size.* The user knows the dimensions of the active area of the tablet.

These three invariants are used in combination with the user's senses of proprioception and kinesthesis (the senses of the position and movement of the limbs) to maintain contextual information throughout the exploration and to be able to point directly to different areas on the data set without having to access information sequentially. Since the complete data set is presented on the tablet, there is no need to scroll information; and because the location of the cells remains static, once some area of interest has been identified on the data set, the user is confident that the data will always be on the same physical position. The tablet has a physical, tangible frame around its active area, which provides easily felt reference to the constant size and location of these boundaries (see Figure 1, b). Feeling these fixed, physical frames, combined with the user's senses of proprioception and kinesthesis, facilitate jumping directly to those areas of interest that have been identified earlier in the data browsing. Furthermore, since the data set fills up the active area on the tablet completely, the user can also jump directly to specific areas on the data set (the top, bottom left or right edges of the data set, the centre, explore the four corners, etc.). Finally, as users traverse data sets obtaining information through sound, they keep track of the position of the pen (of the data under focus) in relation to the physical boundaries of the tablet (that is, in relation to the boundaries of the data set), maintaining constantly contextual information through the uses of sensory modalities other than hearing, which is being used to analyse data. Maintaining focus and context information is very important to support the functional components of browsing, as described by Kwasnik: orientation, place marking, identification, resolution of anomalies, comparison transitions. (Kwasnik 1992).

In addition to several iterations of qualitative evaluations of the interface, two quantitative evaluations have been conducted, both with groups of sighted, blindfolded participants and with groups of participants who were blind or visually impaired. The first study shower that it was faster to obtain overview information from data tables by exploring them using the sonification technique described above, than using a speech synthesizer based interface, while the accuracy of the information obtained and the user's subjective mental workload were similar in both cases (Kildal and Brewster 2005; 2006a).The second study showed that using TableVis to obtain overview information from data tables is highly insensitive to the size of those tables, at least for the range of table sizes that was tested (sizes from 7x4=28 to 31x24=766 cells) (Kildal and Brewster 2006b).

3. EXPLORATORY STRATEGIES AND PROCEDURES IN TABLEVIS

Exploring tabular data sets non-visually is a complex task in which understanding the mathematical relations intrinsic to a tabular structure is used to provide meaning to the data contained in the table, in the context of the metadata of the particular table being explored. In a similar way to exploring a physical object using vision or the sense of touch, a data set in TableVis is explored actively obtaining different views of the data object and combining them to infer properties of that object. The exploratory strategy is the action plan that the user devises to carry out the exploration of a data set in search for specific information. Exploratory procedures are the ways in which that strategy or action plan is implemented; the ways in which actual interaction with the system takes place, including repeatable subroutines that contribute towards building an understanding of the data set. The complexity of the tabular data exploration task and certain degree of flexibility in the interaction with the interface give way to the users' inventiveness to devise their own ways of approaching the problem. Observing these approaches can provide substancial information about which practises should be supported, and it can throw some light on the cognitive processes that take place during the exploration of data tables.

This section examines the exploratory strategies and procedures used by the participants to obtain overview information from data tables, during the evaluation of TableVis. This analysis is done based on data collected through observation and recording of the interaction of the participants with the interface. During the evaluations from which the data analysed here were recorded, the participants had to interact with the system and obtain overview information about numerical data tables. All the data tables had seven columns (days of the week) and twenty-four rows (hours of the day) and the data contained in each cell was the hourly number of visitors that a hypothetical website received at a particular day and time. Thus, a complete table contained information about hourly number of visitors to that website, in a whole week. Two types of questions could be asked about a data table

1. Find which day of the week (or which time of the day) the website is busiest (or least busy).
2. Determine which quadrant of the table contains a cluster of cells with higher or lower values than in the rest of the table.

In the next two sections, data from eight participants observed and recording during the evaluation of the interface are presented and examined. All eight participants were visually impaired (congenitally or adventitiously impaired, with total or partial blindness). From a total of twelve participants who volunteered for this study, only eight of them completed it. Three of the four participants who did not complete the study did not develop an understand the concepts of tabular data structure and the mathematical relations between rows and columns involved in them, nor did they get to understand the data sonification metaphors in the rows and columns modes well enough to be able to complete the session in one hour. They were mature participants (between 35 and 60 years old) who had not studied or utilised mathematical concepts involving tabular structures in recent years. The fourth participant who could not complete the study suffered additionally from hearing impairment that was severe enough not to let him/her discriminate sound pitches. Each participant received training about the interface, the sonification metaphors and the functionality available to browse data in tables. Then, each participant was asked to explore twelve different data sets and obtain overview information to answer to questions of the types i and ii, as shown earlier (six of each type). The order in which these types of questions were asked was counterbalanced. At the end of the evaluation, all the participants who had volunteered where paid for their participation in the study.

3.1. Stages of Exploration

In the observation of the bigger structure of the exploration process of a single table, three different stages could be distinguished. The first stage required mental processing of information by the user, and no action took place. The second stage included all the active interaction with the system, where information from the data set was retrieved interactively. The last stage, mainly or exclusively mental information processing took place again. In the context of the tasks performed by the participants, and in the light of the additional information collected from some of the participants using interview and think aloud evaluation techniques, it is proposed that these three stages can be rationalised as presented in the list below. The description of the stages also includes the

1. *Interpret the question in terms of tabular data structure, and devise strategy for exploration.* The first step requires the user to understand the question and select the best strategy to explore the data in order to find an answer to that question. First, a connection between the question and the metadata of the data set has to be established in order to understand what actions are required next. For example, if the metadata already known by the user establishes that the columns of the table are days of the week and the rows are hours of the day, a question that starts with "at which time of the day…" will be asking about "which row…", and will require a precise row as the answer. Questions of the type *i* required users to explore data tables in rows or in columns modes only, depending on whether it is a row or a column that is being searched for. Questions of type *ii,* instead, required users to scan the data tables both by rows and by columns, and to combine the information obtained from both views to infer the answer to the question. The translation of the question in terms of the meaning of the data into the strategy to explore the data table requires a good understanding of the distribution of data in a table and the relationships between rows and columns. Some of the users that volunteered to take part in the evaluations of TableVis did not complete the evaluation because they could not complete this first stage, due to the fact that they did not have a good enough understanding of the mathematical relations between the cells in a table. This is the stage that arguably makes the highest mental processing demands on the user.

2. *Scan the data in one or two directions.* This stage includes practically all the interaction with the interface. Interactively, all the information is retrieved in sound. The user performs the actions planned in the previous stage, repeating and refining searches until all the information required in terms of tabular data structure and features in the data has been collected.

3. *Combine all the information retrieved through interaction with the system, and obtain the answer to the question.* Having obtained general information in one or two navigation modalities (depending on the task), the user still needs to obtain the answer for the question in terms of the meta-data. If the answer is a particular row or column, the user still needs to interact with the system to request details in speech at the position that had been identified to be containing the correct features in the data. On the contrary, if the task was to identify a region in the table that required exploring the data set in both navigation modes, no further interaction with the interface is required, and the user has to perform the mental exercise of inferring the answer by superimposing both complementary views of the data (obtained in different directions of exploration) and identify a particular area in the table as containing features in the data that are in accordance with the information in each separate view.

3.2. Exploratory Strategies and Procedures

Having described the three stages of exploration, this section analyses the data observed and recording during the second stage, i.e. during the active exploration of the data through interaction with the interface. Figure 2 shows the traces of the pen on the tablet, for each one of the eight participants that completed all twelve data explorations. Each figure in the table includes the traces of all twelve explorations of a participant. The authors also analysed these traces reproducing them as an animation, and observing the exploration of each data set separately. Traces formed with small vertical lines indicate exploration in columns mode. Traces formed with small circles refer to exploration done in rows mode. Lines or circles closer to each other in a trace line indicate that the pen was being moved more slowly on the tablet. Some demographics about the participants, using the same naming convention as in Figure 2: all participants were in the range of 18 to 25 years old, except for participant *d*, who is in the range 46 to 55 years old. Participants *a, c, e* and *g* were totally blind, blindness being congenital for participants *a* and *e*. The other four participants were partially sighted, but required accessibility aids to use computers. Participants *b* and *d* used screen readers, while participants *f* and *h* used screen magnifiers and high-contrast colour schemes.

The differences in the approaches taken by the participants were observed in the exploratory procedures employed during the second stage in which actions are performed to scan the data. They are analysed in the following list:

3.2.1 Use of speed of scan. The speed at which the pen is moved on the tablet, determines the number of sonifications per unit of time that are generated. Thus, faster speeds present mode information per unit of time and are more appropriate to obtain very general information the set of sonifications generated (like trends, patterns, or location of outliers), but worse extract richer information about each one of the sonified rows or columns. The opposite is true for slower speeds. Except for participant b, all the other participants developed a procedure that involved scanning the table by moving the pen at an approximately constant speed, to get an initial overview of the data. Although the speed was constant, it varied substantially between participants. Participant a made use, in average, of some of the fastest hand movement speeds, while participant f used consistently a slow, uniform speed. Participant e made use of the broadest range of uniform speeds between scans, making some of the traverses at fast speed and some others much more slowly. Participant b showed a different approach to the use of speed. This participant scanned the tables slowly, stopping for a short period of time after hearing each one of the sonifications to mentally analyse it. In Figure 2-b, these speed variations can be observed in the frequent variations in density of the trace-plots.

3.2.2 Number of traverses. In most cases, traversing the table only once was enough for all the participants to know the approximate location of the data with interesting features. After the first traverse, they normally went back directly and quickly to the area in which they had identified interested data, simply to confirm the finding. If two or more interesting locations had been identified, they were quickly compared, trying to ignore all the rest of the sonifications that still were being generated as the pen moved. Most participants performed approximately the same number of traverses in all the tables they explored. Participant b rarely required more than one, while participant h rarely performed less than 3 full traverses. All the other participants showed the behaviour described above, i.e. perform a full scan and then confirm the findings locally. As an exception, participant e showed great variability in the number of full traverses performed in each navigation mode, ranging between one and seven. It was also observed that, in general, scans performed at faster speeds of pen movement required more full traverses of the data set, but the whole process did not necessarily take longer time to complete.

3.2.3 Direction of scan. Both in rows and in columns navigation modes, most scans started near the top-left corner of the tables, for the majority of the participants. Clear examples of this are participants d and h, but most of the other participants show this tendency too. An explanation for this choice may lie in the meaning or meta-data of the particular sets being explored. In them, the top-left corner is the origin for both axes. This would suggest that context is maintained during the exploration also regarding the meaning of the data being explored, what would make the final stage of translating the information gathered into a meaningful answer easier. Further investigation is required to confirm this hypothesis. In some other cases where explorations required both rows and columns views, some participants scanned the data very efficiency by drawing lines that followed the areas near the borders, starting a new traverse where the previous one ended, thus not having to lift the pen and relocate it somewhere else to start a new scan. Participants a and c are good examples of this.

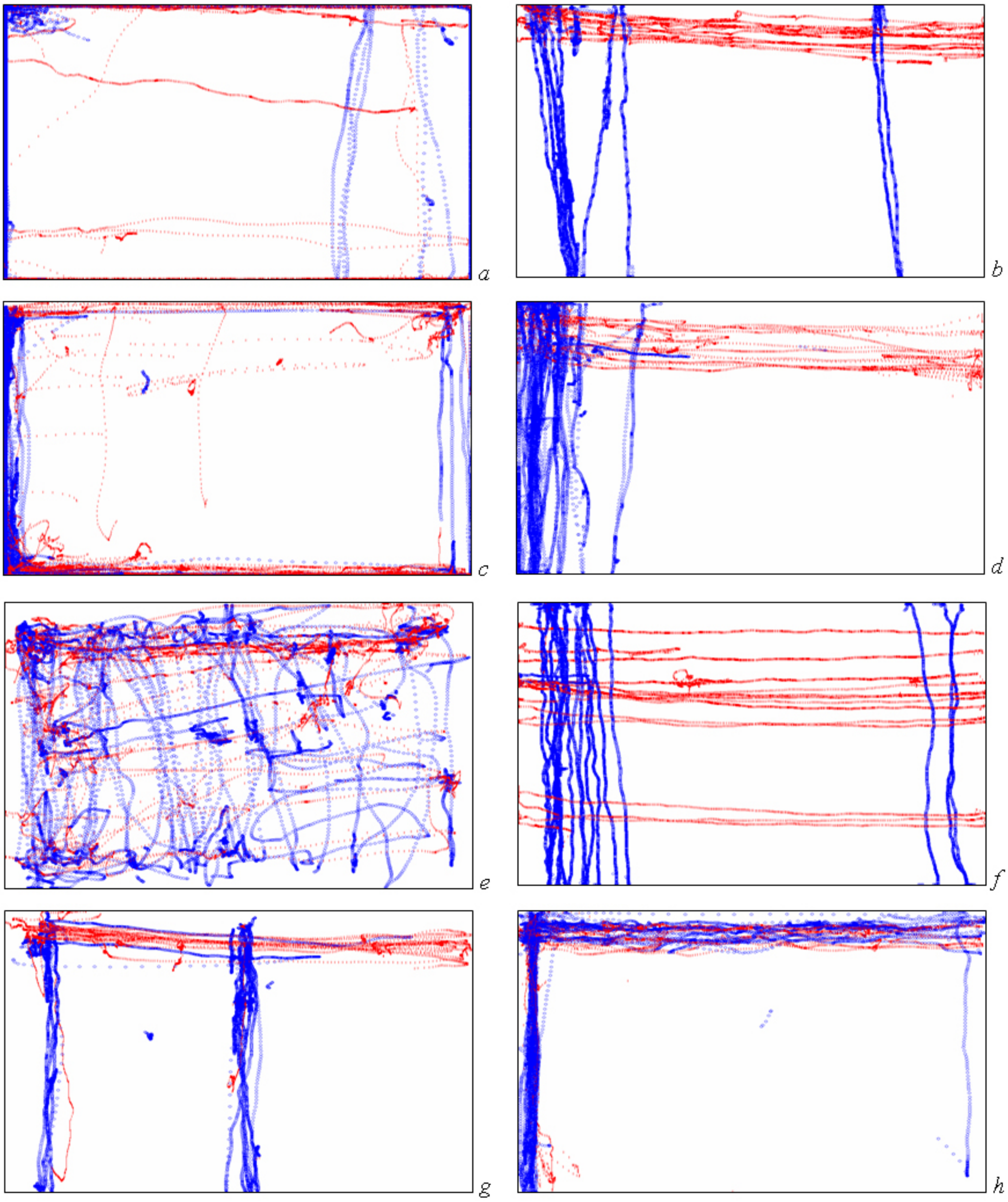


Figure 2. *Traces of the data explorations of eight participants with visual impairments. Each figure contains the twelve explorations performed by a participant. Red or lighter shade traces are generated in columns mode, and blue or darker shade traces generated in rows mode*

3.2.4 Use of tangible, physical borders. The rulers are intended to make the working area tangible to support obtaining context information, and in that sense it is necessary to use them to explore data sets successfully, although this is normally not done consciously. By observing the way in which the tables were scanned and how participants could go directly back to areas of interest, it can be inferred that the rulers were used for this purpose by all the participants. Some participants also chose to use the rulers as guides to traverse the tables horizontally and vertically. Participant a made heavy use of the top and left physical borders, making virtually every exploration in this way. Participant c also used the tangible borders as guides, mainly the left and bottom border, but also the top border. No other participants used the physical borders as guides, except very occasionally. Two considerations can be made here. The first one is that it could be expected that participants with total blindness, particularly those who have a congenital impairment, might find more benefit from using the rulers as guides to traverse the surface of the tablet. In fact, both participants a and c were totally blind. Furthermore, participant c, who used the borders as guides more consistently, was congenitally blind. However, participants e and g (adventitiously totally blind and congenitally totally blind respectively), show that explorations could be performed without any additional support from external physical guides. The second consideration refers again to participant a, who made use of the highest speeds for scanning the data, and also used the tangible frame as guide for the pen most consistently, suggesting the possibility that using the frame in this way can be a technique that may help obtain higher efficiency during exploration.

3.2.5 Shape of lines. Despite the fact that the pen does not need to be dragged in straight lines to scan the table, as log as it travels between left-right and top-bottom, straight lines were the most common shape drawn on the tablet during the data explorations. In addition to the two participants that used the physical borders of the tablet as guides(participants a and c), the rest of the participants, with the exception of e, drew lines that were not only remarkably straight but also very close to being horizontal and vertical. Partially sighted participants f and h are particularly good examples, as it is participant g among the blind participants who did not use the physical borders as guides. Although it could no have been expected that a user kept the pen within a single row or column confidently (and indeed there is no need for it to obtain an overview), the traces presented here show good discrimination between horizontal and vertical, and the ability to stay approximately in a straight line, for participants with a variety of visual impairments. Participant e used a particularly free style of exploration, in which very fast scans were nervously drawn, vaguely following horizontal and vertical lines.

3.2.6 Use of details on demand. With very few exceptions, this functionality was used as it had been intended, that is, only at the end of the exploration and only if detailed information was required in the answer. A good use of this functionality was also made when a participant detected a position during the initial scan that could well be the answer that was being searched. Then, the participant would request details in speech (the text of the label of that row or column) before continuing traversing the table. If at the end of the traversal no other plausible candidates had been found, the answer could already be given without having to go back to that position to get detailed information, because that information had already been collected when it was first identified. The best example of this strategy was followed by participant b, as part of the strategy that this participant followed (see 0 and 0).

3.3. Some common issues and possible solutions

The observation of the exploratory procedures followed by the participants revealed also some common situations that appeared to be confusing or at lest in need of further support from the interface.

3.3.1 Modality confusion. Probably the most common source of temporary confusion emerged when the participant thought that one navigation mode was active while it was in fact the other one that was active, despite the confirmation of the selection received through synthesised speech when the modality was selected. In those cases (for instance, when the participant tried to explore columns with the rows mode active), no sounds emerged from the scanning of the data, except occasionally and erratically (when the pen crossed “by accident” the border between two rows). The severity of this problem is minor, because not obtaining the expected response from the interaction alerted the participants about the problem, which could be rectified immediately. Interestingly, however, it was common for some participants to finish a horizontal scan and start with a vertical one forgetting to change the selection on the navigation mode. Some property in the sonification that was different between both modes could provide constant confirmation about the currently selected modality. Alternatively, having to actively be engaged in the selection of one of the modes could avoid these brief episodes of confusion. Further investigation needs to be conducted about the best ways to solve this problem.

3.3.2 Comparison of non-adjacent rows of columns. A different kind of issue was observed in situations in which a participant had identified two or more possible candidate rows or columns for the answer, which were not adjacent to each other, and comparisons had to be performed. Some participants used the fingers in one hand to mark the locations of those points of interest, and then tried to jump directly between fingers to compare only the sounds that had been pre-selected. Since those positions were not adjacent on the table, moving the pen between them would also produce sonifications of the rows or columns that were in between. This was difficult to avoid even by lifting the pen at one position and putting it down at the next, because the technology used in the wireless graphics tablets (like the WACOM tablet used in TableVis) allows the pen to communicate with the tablet at several millimetres of distance, before they get in contact. For this reason, performing “clean” jumps is difficult. Additionally, marking several points on the tablet with the fingers from one hand is inaccurate and often the reference got lost by having moved the hand. This situation would benefit from some form of external memory aid that could be used to mark interesting positions that could be easily found again, and which also permitted filtering out the information that had not been selected, to perform comparisons between the selected positions without interferences from other sonifications that got triggered in between. Future work will investigate this further

4. CONCLUSIONS

Exploring numerical data tables non-visually to obtain overview information that can answer high level questions about the data, is a complex task. TableVis provides support to complete tasks of this kind through interactive exploration and sonification of the data. Analysing the way in which users with visual impairments approach this problem and make use of the functionality available in TableVis provides valuable insight into the practises that should be supported and the issues that must be tackled.

This paper has focused into observing and examining exploratory strategies and procedures followed by participants with visual impairments that evaluated the interface. A three-stage approach to completing the task has been described. While the general exploratory strategy chosen by all the participants was very similar, the way in which the actual data was gathered varied between participants. A number of exploratory procedures have been described, and further investigation about them is suggested to obtain evidence about which are the best practises and which cognitive aspects of the exploration should be supported better.

Issues regarding discrimination of navigation modes and the need for external memory aids to select and filter information have been identified as well. Future research will address these questions.

Acknowledgements. We want to acknowledge the RNC in Hereford for their committed participation and support of our research studies. We also want to acknowledge the EPSRC for funding the MultiVis project (grant GR/S86150/01) , part of which is the work presented here (www.multivis.org).

5. REFERENCES

Brown, L. M., S. A. Brewster, et al. (2003). Design Guidelines for Audio Presentation of Graphs and Tables. Proceedings of ICAD 2003 Workshop on Auditory Displays In Assistive Technologies, Boston, MA.

Kildal, J. and S. A. Brewster (2005). Explore the matrix: Browsing numerical data tables using sound. ICAD2006, London, Ireland.

Kildal, J. and S. A. Brewster (2006a). Non-visual Overviews of Complex Data Sets. Extended Abstracts of CHI 2006, Montreal, Quèbec, Canada, ACM Press.

Kildal, J. and S. A. Brewster (2006b). Providing a Size-Independent Overview of Non-Visual Tables.

Kramer, G. (1994). An Introduction to Auditory Display. Auditory Display: Sonification, Audification, and Auditory Interfaces. G. Kramer, Addison-Wesley**:** 1-77.

Kwasnik, B. H. (1992). A Descriptive Study of the Functional Components of Browsing. Proceedings of the IFIP TC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction, North-Holland**:** 191-203.

Mansur, D. L. (1985). Graphs in Sound: a Numerical Data Analysis Method for the Blind, California Univ.

Miller, G. A. (1956). “The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information.” The Psychological Review **63**: 81-97.

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. IEEE Symposium on Visual Languages, Boulder, CO, USA, IEEE Comput. Soc. Press.

Walker, B. N. and J. T. Cothran (2003). Sonification Sandbox: a Graphical Toolkit for Auditory Graphs. International Conference on Auditory Display (ICAD), Boston, MA, USA.

Understanding users with reading disabilities or reduced vision: towards a universal design of an auditory, location-aware museum guide

L Pareto and U Lundh Snis

Laboratory for Interaction Technology, University West
P.O Box 795, SE-45126 Uddevalla, SWEDEN

lena.pareto@hv.se, *ulrika.snis@hv.se*

ABSTRACT

We present ongoing work on the design of an information system for users with reading disabilities and users with reduced vision. The design target is a portable, auditory, location-aware information system, to complement visually displayed information in exhibitions. Applying a user-centred, we identify non-typical user-groups' specific requirements, which are turned into a design. The first design-iteration, which includes a formative evaluation, using a mock-up prototype, with dyslectic and visually impaired participants, is completed. The evaluation indicates that the user-group's specific aspects we have identified are relevant, while designing for these groups.

1. INTRODUCTION

Many public places are information intense with several, sometimes large, displays of various kinds, e.g., museums, libraries, train stations, and airports. Information in such environments is not accessible to people with *reading disabilities* (people with minor comprehension disabilities, and dyslexics) or *reduced vision* (elderly people, and people that are visually impaired). According to the International Dyslexia association (International Dyslexia Association, online) 15% of the US population has a reading disability; according to the American Foundation for the Blind (American Foundation for the Blind, online), 2-4 % of the US population is blind or visually impaired. The National board of Statistics in Sweden reports 8-10% of the Swedish population to be diagnosed as dyslexics and 1-2% as visually impaired (Statistics Sweden, online).

This paper takes the point of departure from the design of information systems for these two groups. The two user groups have distinct needs both with respect to *information content* and with respect to the *technical device* providing the information; to design a usable system we must investigate these needs and understand which implication they have on the design. Although their respective needs are different, both groups benefit from having auditory information compensating, augmenting or complementing visual information.

Rather than solely providing assistive technology solutions (Liffick, 2003), the aim ought to be on *universal* design solutions according to Schneiderman (2000) and Law et al, (2005). Mainstream products with universal usability are attractive to a larger population and therefore more cost-effective, an important aspect for most organisations. In alignment with this reasoning, we argue that a suitable strategy for approaching universal design solutions is to identify non-typical user groups and to focus on their respective needs in the design process, without excluding the rest of the population as potential users of the designed system. Hence, we will focus on two non-typical user groups during the design, i.e. people with reading disabilities or reduced vision, respectively, but the aim is to provide auditory complement to visual information for hearing individuals in general.

The case we consider in this work is an auditory museum guide, dedicated to our two user groups. The case study was carried out in collaboration with a regional museum in Sweden, with exhibits reflecting the local culture, local handicraft and artworks. Most exhibits are organized around visual objects accompanied by displays of text. Objects and information is distributed over several rooms and floors. The typical visitor wanders around and experience objects as they appear along the route. The information is distributed (scattered around) and often tied to the location where it occurs. Typically, a person wanders around in the museum; the actual situation triggers the desire for information. Therefore, we are considering portable, context-dependent devices, which provide auditory information to the user at the location, when desired − just as visual displays do. Information systems of this kind need to be location-aware, as well as interactive.

The purpose of this paper is to support the design of systems involving portable, auditory, context dependent information devices to be used by people with reading disabilities or reduced vision. We do this by

- Identifying aspects which are relevant for our respective user groups, and specific to these.
- Eliciting system requirements guided by these aspects for a particular case. (A museum system.)
- Turning these requirements into a conceptual design.
- Evaluating the design with representatives from both our user-groups.

2. UNDERSTANDING THE USER GROUPS

The get a better understanding of the users groups, we undertook the following: We studied literature about dyslexia and visual impairments. We visited a school for adult dyslexics, to perform interviews with the teachers. (To involve knowledgeable persons from communities around a user group—a method discussed in Sullivan and McGrenere (2003)—is often a more effective way to get general information about a group rather than to involve users.) The interviewed teachers' experience, collected during 14 years of teaching, gave valuable information of common problems among dyslexics and insights into methods and equipment used. We interviewed a woman with severe dyslexia who also has a dyslectic son. We contacted the national organization of visually impaired in Sweden (National Association of visually impaired, online), and interviewed the chair of the organization, who also works with rehabilitation of visually impaired people.

2.1 Reading Disabilities

Of the population with reading disabilities, 85% has what is called dyslexia (International Dyslexia Association, online). Dyslexia is a language processing disorder with many different symptoms and combinations of symptoms. Some dyslexics have problems translating language to thought (i.e. reading or listening), others translating thought to language (i.e. writing or talking), yet others have problems with both. Dyslexia is characterized by inefficient information processing. This includes difficulties with phonological processing, the working memory, rapid naming and with automation of basic skills. Contrary to an occasional misconception, it is not a problem of vision: people with dyslexia do not "see backwards". The symptoms for each individual vary, but there are common symptoms related to reading or listening: lack of awareness of sounds in words; difficulties with decoding words,; difficulties with sequences of letters (left-felt); problems with reading comprehension; confusion about directions in space; imprecise or incomplete interpretation of spoken language (International Dyslexia Association, online; Ross-Kidder, online).

Language processing in any form requires an effort for most dyslexics. Many have problems retaining the contextual information while reading or listening. Therefore, material is only comprehensible if it consists of short sentences in short paragraphs. Meta constructs in languages such as metaphors or irony adds a layer of interpretation—as the text cannot be understood literally—further reducing comprehensiveness.

Many dyslexics perceive audio-books and other auditory material as more comprehensible than texts. In auditory material there are distinct pauses between paragraphs, and sometimes relaxing music to indicate and stimulate pause and rest for the listener. Many, however, have problems interpreting auditory material too: specialized tape-recorders with additional controls to change the speed of the narrated text, controls for treble and bass and a button for pause are sometimes used by dyslexics. Since the language processing is demanding and the phonetic interpretation is less clear for many dyslexics, it is harder for dyslexics to mask out the desirable information from the surrounding noise. To some dyslexics, background noise and conflicting sources of sound is very disturbing.

2.2 Reduced Vision

There are many variations of reduced vision, with diverse effects on how the viewer experiences the target. Reduced vision can have effects such as seeing a blurry picture, seeing a grey spot while trying to focus (but clear and sharp around it), seeing only light and dark, or seeing the world through a small camera (reduced area of vision) (National Association of visually impaired, online; Tiresias, online).

Some visually impaired have experienced a normal vision earlier in life, whereas some are born with their impairment. To compensate for a reduced function, humans tend to take advantage of other senses more effectively than others. People with reduced vision can compensate their lack of seeing by developing their listening skills and by using their hands to "see with", for instance. Blind or almost blind people use their hearing to a much greater extent than seeing people do, which indicate that they are trained listeners.

Visual impairment is used as an umbrella term to describe a broad spectrum of individuals and abilities, and particular attributes of a person's disability will make assistive solutions more or less appropriate (Fraser, and Gutwin, 2000). For instance, magnifying is not appropriate for users with limited field of view; change of colour or high contrast is helpful for some types of impairment, but not all; people who retain central vision have other needs than people with limited central vision do. However, complementary auditory information ought to be beneficial for most users with reduced vision, unless they also suffer from reduced hearing. Since many elderly people suffer from both, it is important to make the auditory device compatible with hearing aids.

3. REQUIREMENTS ON AN AUDITORY INFORMATION SYSTEM SUPPORTING BOTH USER GROUPS

Our approach is to study needs and derive requirements from non-typical user groups, as a strategy to achieve more universal, mainstream products. Therefore, we continue by analysing the user groups' respective needs and how these can be combined into one solution. As a result of our user studies, we have identified the following requirements on information systems for users with reading disabilities or reduced vision. The requirements are organized in technology requirements and information requirements.

3.1 Requirements on the Technology:

Since a reading disability often originates in a language processing deficit, functionality that supports these users' language comprehension processes should be provided. The ability to alter the narration speed of speech is important. Due to problems of discriminating sounds within words, dyslexics are often more sensitive to the level of treble and bass, so these parameters ought to be controllable. The possibility to pause is necessary. Possibility to repeat the last sentence (i.e., to back up a semantic unit, rather than a time unit) is desirable. As background noise is an issue for many dyslexics, it should be possible to filter out conflicting sources of sound by, for instance, using isolating headphones.

Users with reduced vision are often accustomed to technical aids, since many depend on these for their everyday activities. Possibility to receive human assistance is considered important among visually impaired. The device's controls should be operable without visual attention: this may be achieved by tactilely distinguishable controls (for instance buttons with relief-symbols), or by auditory feedback when operating the controls. Since users may carry a blind stick or use the other hand for it should be possible to operate the device with one hand only. Users with reduced vision are likely to be accompanied by others, so social interaction is likely to co-occur and must be supported. If headphones are used, they should be compatible with hearing aids, and allow for social interaction.

Differences between the two groups include the ability to hear and comprehend auditory information, and the ability to handle simultaneous sources of sound. This can be solved by providing different alternative sound transmission devices (e.g., headphone, speakers) allowing for private, immersive reception of the information, as well as reception in a more social setting.

Only minimal functionality should be considered when designing for disabled according to Dickinson et al, (2003). Necessary functionality for this system is to provide context-dependent, appropriate information to the user, when the user wants to hear it. The system must provide the possibility to interact with the information (pause, stop, forward, backward, control speed, and perhaps repeat the last sentence, and change treble or bass). The device should be robust, context-sensitive, discrete, and have controls distinguishable with reduced vision (tactilely or with auditory feedback), which is a feature often appreciated even for seeing users. It must allow for listening to information as well as communicating with others and must be compatible with hearing aids. Finally, the system should provide help instructions, and perhaps the possibility to contact human help.

3.2 Requirements on the Information:

In general, the use of simplified language is needed for users with reading disabilities. Avoid difficult and long words as well as complicated grammatical structure. Avoid metaphors, ambiguous use of pronouns and words that are ambiguous by their pronunciation, since they require contextual understanding to interpret. Keep sentences and paragraphs short, to allow frequent rests. Make pauses in between paragraphs longer than normal, and use music to indicate pause. Quality of the audio output is important for dyslectics, in particular the articulation of words. It is preferable to use human, recorded speech, since it is less intelligible and less preferred (Tiresias, online) and easier to perceive for dyslexics than synthetic speech (Blomert and Mitterer, 2004).

The purpose of auditory information for users with reduced vision is, in addition to provide information, to convey environmental aspects (e.g., description of the room, light circumstances, placement of objects in the room) as well as navigational aid. Visual aspects of objects of interest such as size, colour, and main impression, should be explained. A suggested principle is to always explain the "big picture" first, and continue with details after. It is important to avoid referring to visual impressions without explaining them (which seeing people tend to do all the time). To use vivid, metaphorical descriptions is a natural way to explain visual sensations; for this user group, such descriptions are preferable to plain descriptions.

The two user groups have different information needs, regarding the type, amount, and level of information, e.g., information to support navigation is relevant for users with reduced vision only. Users with reduced vision typically appreciate colourful, metaphorical descriptions, whereas dyslexics typically find such descriptions incomprehensible. A strongly simplified language may be considered indigent by the visually impaired, who wants to compensate their lack of vision by descriptive information. Thus, the information must be adapted to each group, respectively.

4. DESIGN APPROACH

In our search for design solutions, we have not found any existing system fulfilling all requirements above. The location-aware, auditory handheld museum guide in Ciavarella and Paternò (2004) is ruled out since the user must choose the object of interest from a visual display on the PDA which is not appropriate for visually impaired people. The Audio-Augmented Reality prototype in Bederson (1995) is a location-aware, fully automated digital tape-recorder, which provides information in the visitors' headphones as soon as the user approaches an object. According to Dickinson et al. (2003) this is in conflict with the need of our users: automated actions tend to be confusing and intrusive to the users, so user-initiated and user-controlled systems are more desirable (Aoki and Woodruff, 2000). The audio-based, location aware system, Ec(h)o (Hatala et al., 2004) are based on general user models and is therefore not suitable for non typical users. Other systems are more concerned with augmenting the museum experience by adding features rather than striving for universal usability, for instance the system in Terrenghi and Zimmermann (2004) provides personalized sound illustrations based on movement and gestures, and the Discovery Point system (Berkovich et al., 2003) encourage social interaction between visitors by initiating short stories about the artefacts in front of the visitor. Thus, we see a need for a new system to be designed differently.

To design technology for users with reduced functionality is challenging for at least two reasons: First, for most designers, it is more difficult to understand these users' situations and needs, since most designers have little own experience and knowledge of disabilities. As described in Dickinson et al (2003): "Developing systems for "non-typical" groups meant that in some cases a "logical interface" in the developer's sense of the term was not a solution." And thus, secondly, standard solutions are not always appropriate, which calls for creative, new solutions which must be tested by representative users. Hence, a user-centred, iterative approach to design which involves users is crucial when designing for disabled. This conclusion is also supported by Dickinson et al (2002), Dickinson et al (2003), and Wattenberg (2004).

The overall methodology we used is a user-centred approach to interaction design (Dix et al, 1998; Preece et al, 2002), in a contextual setting (Beyer and Holzblatt, 1998). So far, we have completed the first cycle of iterations: analysis, design, prototyping and evaluation. In this first iteration of the design work, we have conducted user and task analysis, set up requirements and design implications, constructed a mock-up prototype simulating the system, and performed user tests in the museum with potential users from the target groups.

To evaluate the conceptual design, two sets of user tests were conducted: one with each target group. The test was performed in the museum, in a particular exhibition of paintings, which was the preferred choice of the museum for this study. For the purpose of these tests, a mock-up prototype was constructed, which simulated the device and the main functionality of the system. Auditory information related to 4-5 paintings was developed, following directions from our user studies. The evaluation consisted of (1) users testing the mock-up prototype according to an arranged scenario in the chosen exhibition, (2) observations and logging of their interaction and their reactions when interacting with the device, (3) interviews with the participants directly after the test, and (4) analysis of the texts carried our by teachers from a school of adult dyslexics. The group of dyslexics consisted of 7 and the group of visually impaired of 5 participants. Test groups sizes were determined on the basis of the tests' purpose: the tests were intended as formative evaluations to get input to further design. Such tests motivate small test groups (Nielsen and Landauer, 1993). In accordance with Dickinson et al (2003), we will perform tests continuously during the development, to assure the usability of the device and the content.

5. DESIGN SOLUTION

As a consequence of the different needs regarding information content of the studied user groups, the design solution is split in two parts: a generic infrastructure consisting of a location-aware, context-dependent technology, and specialized content which provide the user-groups with adapted information, respectively.

5.1 Infrastructure: Location-Aware Technology

The infrastructure consists of a portable device together with a headset for auditory information, and an environment in which objects and information displays are equipped with identification tags. The location-awareness in our prototype is provided by IR (InfraRed light) tags on objects, an IR transmitter in the device, and a database with object-related information. The RFID (Radio Frequency Identification) technology (Das Raghu, 2004) was considered, but it turned out to have too many technical limitations for now. The current IR solution is flexible as new tags can be added to the environment; new information can be added or changed in the database and since the context-dependent information is any auditory information. The same infrastructure can be used for various applications. The system includes a hand-held device to interact with, and different kinds of headsets to support different needs: isolating headphones, ear plugs, or small speakers for group interaction.

5.2 Content: User-Group Adapted information

The information is organized in a database, coupling the pieces of information to the identification of the corresponding objects or displays. Content must be provided and adapted to the target groups. For this purpose, guidelines of adaptation to the user groups are provided. The content can either consist of sound files to be transmitted directly, or by text that is transformed into synthesized speech. For the first evaluation with the mock-up prototype, example texts were produced by the museum and recorded in our studio. For each painting, the museum produced information adapted to visually impaired and to dyslexics, following the directions from our analysis. This procedure was time consuming, and is not reasonable in the future. Therefore, to be practical in a commercial application, software for producing content must also be provided.

6. EVALUATION OF THE CONCEPTUAL DESIGN

We have conducted two sessions with formative evaluations: one with each user group. For this purpose, a mock-up prototype simulating the system was constructed. The evaluations consisted of a test with the mock-up prototype, observations during these tests, followed by interviews with the participants. The tests were performed in the museum context, in a dedicated room with paintings. General information about 5 paintings was available, in two versions: one for the group of dyslexics; one for the group of visually impaired. The purpose of the user tests was to evaluate the functionality of the system, the interaction with the (mock-up) device, the appropriateness of the content, the quality of the sound, and the overall experience as perceived by the participants.

7.1 The Mock-up Prototype

Our first mock-up prototype was a simulation of the system and its main functionality. It consisted of a device, two wireless headphones, a laptop, and an operator simulating the system's behaviour. The device we used was a so called Daisy-player (see Figure 1)—a special device developed for dyslexics. The Daisy-player was suitable as mock-up device, since it has all the functionality and controls that are considered for the design. Moreover, there is an option to have the device give auditory feedback when a control is used (such as "Lower the volume"). This feed-back was sent wirelessly to the operator's head-phones, so that the operator could act accordingly. Similarly, auditory information was sent wirelessly to the test user's headphones from the laptop according to the operators commands. On the laptop, an mp3 player was playing the files according accordingly. Test user's could, thus, move around freely, operate the device, and receive auditory information in headphones in the very same way as the system is going to function (see Figure 2).

Interaction commands are sent wireless to operator's headphones Laptop sends sound information wireless to user's headphones Our first attempt to balance user- and system-initiated transmissions in the prototype was to alert the user by a sound, but then expect the user to initiate information playback by pressing "play". This was simulated by the operator playing a special sound when the user came close to objects with available information. In this way, the user was made aware of the presence of information, and was allowed choose to listen to it at will.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

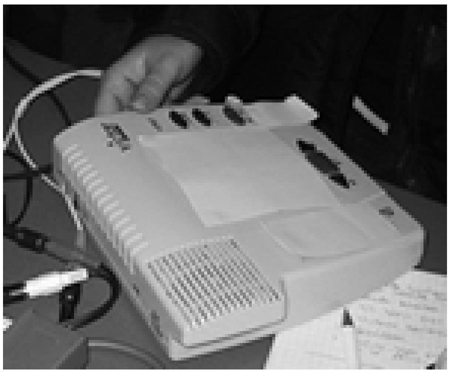


Figure 1. *Mockup device.*

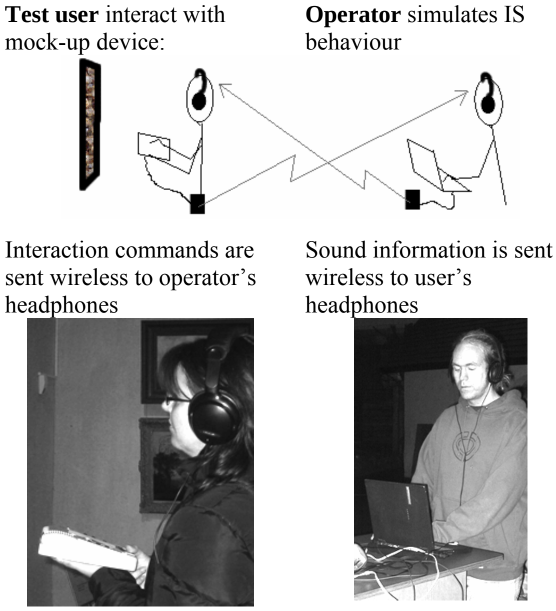


Figure 2. *Prototype simulating the system.*

Figure 3. *Test situation.*

Figure 4. *Test users.*

7.2 The User Tests

The tests proceeded as follows: The participants were introduced to the mock-up prototype, and we explained which functionality it had and how to operate it. They were told to walk around in the room, looking at the paintings as they wished, and trying to act normally (see Figure 3). We carefully explained that the purpose of the test was to test the system, not to test individuals; this was essential, as people in these user groups are often subject to tests considering individuals. We encouraged our participants to be as critical as possible to the system. While they were wandering around listening to the information and handling the device, we tried to observe what they did with the device and how they reacted to the response they got. That information was used in the interview following the test, as points of discussion.

7.2.1 Test users with Reading Disabilities. The first test session was performed with dyslectic people. There were 7 participants, with various degrees and symptoms of dyslexia; all were women of age 20-35. Their interest for art varied; none was a frequent visitor of the museum, but some had been there before. The results from the tests with dyslexics were as follows:

Technology: For most, interacting with the device was unproblematic. However, the use of the Daisy-player as a mock-up device may have affected this result, since the participants were already accustomed to it. Most test users tried to interact with the information by pausing, rewinding, forwarding and stopping. The way the backward button was handled in the prototype, did not correspond to all users anticipation. Some expected it to rewind (as a tape recorder) when hold and others pushed it several times to achieve the same effect. The function to get to the start of a text was a function that several users wished to have. A few adjusted the volume; none adjusted the treble and bass. All participants appreciated the isolated headphones; none considered it a problem to wear them.

Information: The text adapted to the dyslectic participants, was considered reasonable, but too simplified: more colourful descriptions were asked for. After the tests, the teachers who we had interview, analyzed and commented the texts that we had produced. The quality of sound was considered good, but parts of the reading were considered too monotone, even though the texts were read by a professional museum guide.

The overall impression of the dyslectic test group was that they liked the idea: at least they really appreciated that they were put in focus for the development. Some of them said they would consider visiting the museum (which most of them had not done before) and use such a system if it was available.

7.2.2 Test users with Reduced Vision. The second session of the evaluation was performed with visually impaired people. There were 5 participants: all elderly women (see Figure 4). Their degree of vision varied: some could only identify light and dark, and they had various kind of visual impairment; one could only see parts of a painting simultaneously, one could not see in the focus spot and the others had general reduced vision. Their interest of art varied, but several of the participants used to be frequent visitors to this museum; they had, however, stopped visiting as their ability to see was too restricted for a visit to be meaningful. They were all used to various kinds of technological equipments, since they all depended on that for their everyday activities. The results from the tests with visually impaired were as follows:

Technology: The device was a bit difficult to handle, for most of the participants. This can partly be explained by that the buttons did not have enough tactile information on them to be able to identify easily. The buttons also needed to be more visually distinct from the rest of the device. Most of the test users were not as eager to explore the functionality of the device, so after listening to the information, they just continued. However, several participants said they wanted to repeat the description, or parts of it, and listen to the information again. All participants chose and appreciated the isolating headphones, even though smaller alternatives were present. The relation between the stop, start and pause buttons and the actions were not clear (in particular, the difference between stop and pause was not apparent).

Information: The adapted information was much appreciated; they considered the descriptions to be involving and interesting. However, most of the participants wanted more information of the same kind: about the artist, for instance. The quality of sound was considered good. There were requests of more instructions regarding navigation, such as obstacles or where places to sit and rest were to be found. Instructions of available choices, was asked for. On-line help instructions was not provided in the test, only manual description before the test, which apparently was not enough.

The over-all impression of the visually impaired was positive: 3 out of 5 said that they gladly would visit the museum again, if such a device was available to them.

7. CONCLUSIONS AND FUTURE WORK

The evaluation gave valuable input for the next iteration of the design process—which was its main purpose. None of the design aspects we identified have shown to be unnecessary, but some may be of less importance such as the ability to alter treble and bas. The concrete device and the interaction with the device must be specified in detail, prototyped, and evaluated in test with users again.

The two user groups—users with reading disabilities and users with reduced vision—have been in focus throughout the development, and will remain in focus. However, the information system can be beneficial for any museum visitor, since non-typical user groups often have higher, but not necessarily different demands than others. Hence, to focus on several non-typical users groups during design is a good strategy to aim for main stream products with universal usability (Shneiderman, 2000; Law et al, 2005), rather than particular assistive technology (Liffick, 2003).

Future work includes developing and testing more and diverse types of content; testing the first running IR-based prototype, with more and different user groups; experimenting with sound manipulation and sound recording, and developing practical support for content production.

8. REFERENCES

American Foundation for the Blind. Available at http://www.afb.org/

Aoki, P. and Woodruff, A., (2000). Improving electronic guidebook interfaces using a task-oriented design approach, In DIS '00: Proceedings of the conference on Designing interactive systems, p 319—325, New York City, NY 2000. ACM Press.

Bederson, B. (1995). "Audio Augmented Reality: A Prototype Automated Tour Guide," ACM SIGCHI '95 Conference Companion, Denver, CO, pp. 210-211.

Berkovich, M., et al. (2003). Discovery point: enhancing the museum experience with technology, In CHI '03 extended abstracts on Human factors in computing systems, p 994-995, Ft. Lauderdale, Florida, USA 2003. ACM Press.

Beyer, H. and Holzblatt, K. (1998). Contextual Design: Defining Customer-Centered Systems. San Fransisco: Morgan Kauffman.

Blomert L. and Mitterer H. (2004). The fragile nature of the speech-perception deficit in dyslexia: natural vs synthetic speech. In Brain and Language, Volume 89, Issue 1, April 2004, P 21-26.

Ciavarella, C. and Paternò, F. (2004). "The design of a handheld, location-aware guide for indoor environments", Personal Ubiquitous Comput, vol. 8, no 2, 82-91. Springer-Verlag.

Das, Raghu, (2004). An Introduction to RFID and Tagging Technologies, IdTechEx Lmt, Cambridge UK, online: www.idtechex.com

Dickinson, A., Eisma, R. and Gregor, P. (2003). Challenging interfaces/redesigning users, In Proceedings of the 2003 conference on Universal usability, P 61-68, Vancouver, British Columbia, Canada 2003. ACM Press.

Dickinson, A., Gregor, P. and Newell, A. (2002). Ongoing investigation of the ways in which some of the problems encountered by some dyslexics can be alleviated using computer techniques, In Proceedings of the fifth international ACM conference on Assistive technologies, P 97-103, Edinburgh, Scotland 2002. ACM Press.

Dix, A., Abowd, G., Beale, R. and Finlay, J. (1998), Human-Computer Interaction, Prentice Hall Europe.

Fraser, J. and Gutwin, C. (2000). A framework of assistive pointers for low vision users, In Proceedings of the fourth international ACM conference on Assistive technologies, p 9-16. Arlington, Virginia, US. ACM Press.

Hatala, M., et al. (2004). Ontology and rule based retrieval of sound objects in augmented audio reality system for museum visitors, In Proceedings of the 2004 ACM symposium on Applied computing, p 1045-1050, Nicosia, Cyprus 2004. ACM Press.

International Dyslexia Association, Available at http://www.interdys.org/

Law, C., Jacko, J., Peterson, B., and Tobias, J. (2005): Universal designs versus assistive technologies: research agendas and practical applications. In *Proc of the 7th international ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, MD, USA, October 2005). Assets '05. ACM Press, NY, 2-3.

Liffick, B. W. (2003). Assistive technology as an HCI topic. *J. Comput. Small Coll.* 19, 2 (Dec. 2003), 142-144.

National Association of visually impaired, Sweden. Available at http://www.srfriks.org

Nielsen, Jakob, and Landauer, Thomas K. (1993). A mathematical model of the finding of usability problems, in Proceedings of ACM INTERCHI '93 Conference (Amsterdam, The Netherlands, 24-29 April 1993), pp. 206-213.

Preece, J., Rogers, Y, and, Sharp, H (2002). Interaction Design – beyond human computer interaction. Wileys, 2002.

Ross-Kidder, K. Ph.D. Director of Learning Disabilities Online, http://www.ldonline.org/

Shaywitz S., et al. Persistence of Dyslexia: The Connecticut Longitudinal Study at Adolescence, In PEDIATRICS Vol. 104 No. 6 December 1999, pp. 1351-1359.

Shneiderman, B. (2000). "Universal usability", Communications ACM, vol. 43, no 5, 84-91. ACM Press.

Statistics Sweden. National board of statistics in Sweden. Available at http://www.scb.se/

Sullivan, J. and McGrenere, J. (2003), Designing cognitive technologies for people with disabilities - perspectives for theory and practice, In Proceedings of the 2003 conference on Universal usability, p 148-149. ACM Press.

Terrenghi, L. & Zimmermann, A. (2004).Tailored Audio Augmented Environments for Museums, In Proceedings of the 9th international conference on Intelligent user interface, Funchal, Madeira 2004. p. 1006-1013, ACM Press.

Tiresias, International Information on Visual Disability. Guidelines of Audio Output for Visually Impaired. Available at http://www.tiresias.org/.

Wattenberg, Ted, (2004). Beyond standards: reaching usability goals through user participation, In SIGACCESS Accessibility and Computing. No 79, p 10-20. ACM Press

Preliminary work for vocal and haptic navigation software for blind sailors

M Simonnet, J-Y Guinard and J Tisseau

European Center for Virtual Reality (C.E.R.V.), École Nationale D'Ingénieurs de Brest, FRANCE

mathieu.simonnet@univ-brest.fr, jean-yves.guinard@univ-brest.fr, tisseau@enib.fr

www.pageperso.univ-brest.fr/~simonnet/anglais.htm, www. cerv.fr/~tisseau

ABSTRACT

This study aims at the conception of haptic and vocal navigation software that permits blind sailors to create and simulate ship itineraries. This question implies a problematic about the haptic strategies used by blind people in order to build their space representation when using maps. According to current theories, people without vision are able to construct cognitive maps of their environment but the lack of sight tends to lead them to build egocentric and sequential mental pictures of space. Nevertheless, exocentric and unified representations are more efficient (Piaget et al, 1948). Can blind people be helped to construct more effective spatial pictures? Some previous works have shown that strategies are the most important factors in spatial performance in large-scale space (Tellevik, 1992) (Hill et al, 1993) (Thinus-Blanc et al, 1997). In order to encode space in an efficient way, we made our subject use the cardinal points reference in small-scale space. During our case study, a compass establishes a frame of external cues. In this respect, we support the assumption that training based on systematic exocentric reference helps blind subjects to build unified space. At the same time, this training has led the blind sailor to change his haptic strategies in order to explore tactile maps and perform better. This seems to modify his processing of space representation. Eventually, we would like to study the transfer between map representation and environment mobility. Our final point is about using strategy based on cardinal points and haptic virtual reality technologies in order to help the blind improve their spatial cognition.

1. INTRODUCTION

This paper describes how we are taking into account blind people's spatial cognition in order to conceive simulation software for blind people's navigation on a virtual sea.

Different spatial theories accord various spatial capacities to blind people (Fletcher, 1980). However, we know that "the main characteristic of spatial representations is that they involve the use of reference (p.11)" (Millar, 1994). The lack of sight tends to lead to a body centered spatial frame, but maps are external reference frames (O'Keefe et al, 1978). How can we make exocentric reference easier for blind people when they encode space?

"Search strategies in haptic exploration are related to encoding processes (p.223)" (Tellevik, 1992). Therefore, it means that teaching blind people exocentric representation should help them to use haptic exploration strategies more effectively. Exocentric strategies aim at locating object to object regardless of the person's body position. This top-down reasoning appeals to cognitive level in the first place and sensory-motor level in the second place.

The north is the prior knowledge acquired by map-readers (Lloyd, 2000)) so we taught blind people that the concept of cardinal points is an absolute exocentric reference. Thus, we made them practise reasoning about the cardinal concept in order to build exocentric spatial representations. According to Vygotsky's socio-cultural theory (1896-1934), psychological instruments like writing and maps are able to reorganize individual cognition. The compass is one of them. After testing our spatial reference task, we have experimented systematical training with a compass. The analysis of these results focuses on the haptic exploration strategies.

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

In the future, we shall consider the essential question of the interest of cardinal strategy for transferring spatial capacities between maps and environment. Our position is that virtual reality can help blind people to connect micro and macro scales in the same exocentric reference frame.

2. NON VISUAL SPATIAL THEORIES

After a survey of debates about non-visual representation during the last century, we are going to emphasize the distinction between egocentric and exocentric reference frames related to the lack of vision.

2.1 History

How do blind people build efficient space representations? During the previous century different theories tried to answer this question and many controversies appeared about the role of previous visual experience.

According to the inefficiency theory (Revesz, 1933), blind people are able to build unified space representations only from simple forms or elements. The results of a wooden blocks recognizing task show that “touch alone is not as efficient in the perception of (…) complex tactual form relationships as touch aided by visual images (p.13)” (Worchel, 1951). The results of a second experiment about direction estimations in a triangle completion task leads to conclude that kinesthetic cues were better able to perform when translated into visual images (Worchel, 1951).

The difference and inefficiency theories disagree. Fletcher assumes that previous results come from a testing artifact (5). Other experiments show no difference between blindfolded and congenitally blind adults when they were dealing with material that was not “optically familiar” (Juurmaa, 1965). Eventually, “lack of vision slows down ontogenic spatial development…but does not prohibit it” (Kitchin et al, 1997).

To conclude, we have to remember that during a spatial inference task, congenitally blind people performed as well as blindfolded (Rieser et al, 1986). So, the congenitally blind persons are able to construct spatial cognitive maps but this capacity develops more slowly. The question is how the lack of vision slows down the construction of exocentric representation.

2.2 From egocentric to exocentric reference

“Exceptions notwithstanding, there is general understanding that in an egocentric reference frame, locations are represented with respect to the particular perspective of a perceiver, whereas an allocentric reference frame locates points within a framework external to the holder of the representation and independent of his or her position (Klatzky et al, 1998)”.

Vision is the first perceptive modality of data concerning spatial environment (Hatwell, 2004). It gives simultaneously varied information about objects and their configuration in distant space. In addition to visual modality, haptic or tactile-kinesthetic modality informs well about spatial layout (Hatwell et al, 2000). The sequential characteristic of haptic modality leads blind people to encode an environment in successive reference to their own body before executing spatial inference between external objects. Nevertheless, vision is neither necessary nor sufficient for spatial coding (Millar, 1994).

According to the Piagetian theory (Piaget et al, 1948), during their development children construct external reference frames from egocentric reference in coordinating vision and tactile perception. When using haptic modality, blind people have to remember the different segments of the object as a whole in short-term memory. This cognitive effort allows them to construct unified exocentric representation of space (Avraamides et al, 2004). In this respect, a recent experiment shows that “allocentric (exocentric) relations can be accurately reported in all modalities […]” (Carreiras et al, 1992). Thus, spatial representation is not limited to any particular sensory modality although processing is probably faster with vision.

To conclude, not all blind people really build an external frame of reference but they are able to do it. Differences in coding strategies are more implicated than their capacities of spatial perception (Millar, 1994).

3. HAPTIC EXPLORATION STRATEGIES

Since the 1990s, researchers have correlated exploration patterns with the nature of non-visual spatial representation. Evidence of exocentric reference superiority leads us to use semantic representation in order to help blind people to improve encoding processes. We will present our hypothesis about the role of the compass on spatial representation.

3.1 The Known strategies

Tellevik first tested three patterns of non-visual exploration. In this task, blindfolded subjects had to find objects in a large-scale environment (Tellevik, 1992). Using “perimeter” patterns, subjects explored the boundaries of given area. In “gridline” patterns, the subjects investigated internal elements of the area to learn their spatial relationship. With using “reference-point” patterns, subjects relate their exploration to salient elements. The results show that search strategy in haptic exploration may be differentially related to encoding processes. With “perimeter” and “gridline” patterns it was more difficult for the blind to change their perspective than with “reference-point” strategies. This shows that the latter pattern is more exocentric than the others. So, we think “gridline” patterns do not really give information about the relation between elements. Consequently, “gridline” pattern is an egocentric strategy.

One year later, Hill emphasized a lack in literature about object-to-object relationships (Hill et al, 1993). In a direction estimation task about explored environment the results showed that “perimeter” pattern is a “self-to-object” strategy. Furthermore, an “object-to-object” pattern is identified and linked to distance between object reasoning. Eventually an efficient chronology of these patterns seems to involve “perimeter” and “object-to-object” strategies.

In this respect, Thinus-Blanc (1997) studied the correlation between exploration patterns and spatial performance in locomotion and handling space. Subjects without vision have to detect the changes in a previously explored spatial layout. The same two types of patterns of explorations are found in small and large-scale space. On the one hand, “cyclic patterns consist in visiting a sequence of objects, with the same one beginning and ending the cycle (p.36)”. On the other, “the back-and-forth pattern is characterized by repeated trajectories between two places (p.36)” (Thinus-Blanc et al, 1997). In accordance with the O’Keefe and Nadel theory, similarity between the first type of strategy and route knowledge and the second type and map knowledge let us emphasize that “cyclic” pattern is an egocentric reference frame and “back-and-forth” pattern is an exocentric frame of reference (O’Keefe et al, 1978). Here the results verify the superiority of exocentric reference frame.

To summarize research, Ungar (2000) carried out a literature survey of cognitive mapping without visual experience. A synthesis of non-visual exploration patterns identifies seven distinct exploration strategies or patterns: home base-to-object, perimeter, grid, cyclic, perimeter-to-object, back-and-forth and object-to-object strategies which are summarized in table 1.

Recently, a doctorate thesis about “perception and cognition of space by individuals who are blind or have low vision” introduced a new strategy called “perimeter-to-center” (Schinazi, 2005). The subjects explored a constructed maze, located and remembered the positions of six different salient points. The results emphasize two egocentric strategies: “grid” and “perimeter-to-center”. The first strategy consists in exploring the boundaries to identify the shape, size and key features of the area around the perimeter and then the inside of it. We have added this egocentric strategy in table 1.

Table 1: *Ungar (2000) modified table: Nature of strategies identified in the studies by Hill, et al. (1993), Thinus-Blanc (1997) and Schinazi (2005).*

Strategy	Description	Nature
Home base-to-object (Hill et al, 1993)	Moving repeatedly between the home base (origin point for exploration) and all the others in turn	Egocentric
Perimeter (Hill et al, 1993)	Explored the boundaries of an area to identify the area’s shape, size and key features around its perimeter, by walking along the edge of the layout	Egocentric
Grid (Hill et al, 1993)	Investigated the internal elements of an area to learn their spatial relationships, by taking straight-line paths from one side of the layout to the other.	Egocentric
Cyclic (Thinus-Blanc, 1997)	Each of the four objects visited in turn, and then returning to the first object	Egocentric
Perimeter to center (Schinazi, 2005)	explored the boundaries to identify the area’s shape, size and key features around the perimeter and then inside of it	Egocentric
Perimeter to object (Hill et al, 1993)	Moving repeatedly between an object and the perimeter	**Exocentric**
Back-and-forth (Thinus-Blanc, 1997))	Moving repeatedly between two objects	**Exocentric**
Object to object (Hill et al, 1993)	Moving repeatedly from one object to another, or feeling the relationship between objects using hand or cane.	**Exocentric**

All these strategies come from movement observations. The results show that blind people performed better when they used exocentric patterns. This evidence proves the positive correlation between a higher cognitive spatial level and exocentric strategies.

3.1 The cardinal strategy: a top-down process

A coherent relationship between mental representation and sensorial information provides a semantic encoding. Thus, the exocentric or egocentric nature of previous spatial strategies results from cognitive processes. The top-down process, from the map concept to environment stimuli and the bottom-up process tightly fit into each other.

Subjects without vision have the same stimuli at their disposal. As they do not similarly perform, they probably do not use the same mental space concept. How can we induce blind people to use maps as representations?

Cognitive mapping processing requires external cues in long-term memory. We know that "the fact that the information which is reliably available in long-term prior experience influences modes of coding explains coding in blind conditions (p.153)" (Millar, 1994). As we have already seen, the cognition of the north is one of the key prerequisite in order to read a map.

Why not teach blind people cardinal points concept?

Acredolo et al. (1975) explain, "information related to the immediate goal of an action is remembered more effectively than is information that is not (p.221)" (Tellevik, 1992). In this respect, we ask our blind subject to remember spatial layout using cardinal points reference. This learning requires the use of a tactile compass in order to provide salient external cues. We conducted an exploratory experimentation in order to evaluate the efficiency of cardinal reference in space encoding.

4. EXPERIMENT

As we have already seen egocentric and exocentric spatial representations exist. This experiment attempts to observe if the compass leads to the use of haptic exocentric patterns of exploration. At the moment the subject of our exploratory experimentation is an adventitiously blind individual. The man who is forty-five years old, lost his sight at twenty-two. He agreed to be the first to test the following protocol.

4.1 Reference Task

The spatial task consists in reproducing a small-scale spatial layout in an absolute reference after changing position around a table.

4.1.1 Situation. Three square metal sheets are placed at three different points, 90° rotated around the table. The sheets are twenty-five centimetres wide. We use six magnetic pieces of various geometric shapes such as a triangle, a cross, a trapezium, a disk, a half-disk and a square. Each piece is covered with different textures: soft, rough, wire netting, cardboard, tactile lines and crossed tactile lines. All this aims at helping the subject to distinguish all these different objects from each other. The pieces are placed only on the first sheet; on the other two they are placed next to them. A non-tactile gridline is drawn on this handling space. Each magnetic piece is placed in the middle of a five centimetre non-tactile square. The grid lines allow us to measure errors when the blind subject reproduces the layout (*cf.* figure 1).

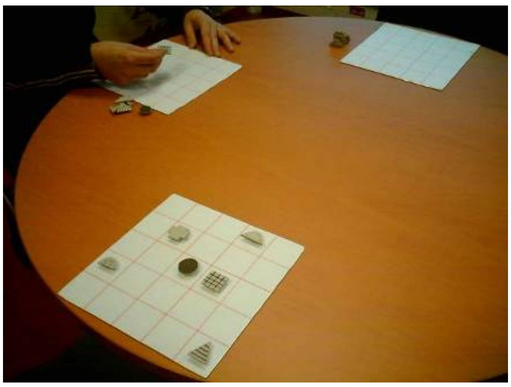


Figure 1. *Reference task illustration.*

4.1.2 The subject experimental activity. The subject sits at the round table and listens to the instructions. After haptically exploring the layout of the six elements on the first sheet without any time limit, he has to reproduce the first configuration on the two other empty sheets. The main point is that this spatial layout reproduction has to be in reference to absolute space and not to body position. So after exploring the first board with the pieces, the subject rotates 90° round the table and manually reproduces the configuration on the second board. He does the task twice.

4.1.3 Collect of results. Since the beginning the subject knows that results depend on the correct positioning of the pieces on the grid drawn on the sheet. The further the magnetic piece is situated away from the correct position, the more important mistake**.** On the one hand, in order to observe the subject's haptic exploration strategies, the tasks are videotaped. Visualization allows us to identify the different haptic exploratory patterns the subject uses. On the other, in order to try to study the cognition of the subject, he is asked to verbalize his reasoning.

The interpretation of this experiment consists in comparing the differences between performance before and after learning the cardinal points. This aims at evaluating the impact of our cardinal strategy in training.

4.2 Training Tasks

The cardinal training consists of three training sessions. All the while the subject could use a tactile compass. We continue the training until the subject is successful.

4.2.1 Task one: Cardinal Orientation question. The cardinal orientation question task is composed of two parts. The instructor places one magnetic piece and asks him to tell the relevant cardinal orientation between the piece and the centre of the sheet. Then the instructor asks questions about the cardinal orientation between two objects placed randomly on the table. After each answer, the subject is given feedback by the instructor. The answers can be north, south, east, and west; northwest, southeast or north-northwest, east-southeast… After one correct answer, the subject stands up and walks a 90° rotation before sitting down in front of the next sheet. The instructor questions him about the cardinal orientation of another magnetic piece and stops after three corrects consecutive answers. At the end of this task, we can assume that the subject has internalised cardinal map representation.

4.2.2 Task two: Cardinal Orientation positionning. This task consists in positioning elements around the center according to the instructor's request. In the second part of the task, the subject is asked to place two objects on a cardinal axis such as southeast northwest for example. At the end of this task, we assume that the subject is able to apply his cardinal map representation to the physical environment. Thus the subject can now make use of cardinal orientation positioning.

4.2.3 Task three: spatial layout production. The subject puts the six pieces wherever he likes on one sheet. Afterwards he has to do it again on the other two sheets. In this final task, we attempt to enable the subject to get into the habit of building his own favorite constants in the exocentric cardinal reference frame. For example, using the northwesterly corner as a reference point seems to be efficient.

When the previous three tasks are successfully completed, we have to wait one week before asking our subject to perform the reference task anew in order to avoid the straight recall effect of learning (Schmidt, 1975).

5. RESULTS

Before cardinal training, the subject had made seven mistakes. From a strategy point of view, on the one hand we clearly identified "home-base-to-object", "cyclic" and "grid" egocentric patterns of haptic exploration and on the other, the "perimeter-to-object" exocentric strategy appears. In other words, the subject uses mostly egocentric spatial representation**.**

Some other behavioral cues emphasized this assumption. The video recording showed egocentric behaviour during the spatial layout reproduction task before cardinal learning. The subject tried to turn the board in front of him before performing. As this was not allowed, he first used body references and put the pieces in wrong squares and then changed their position along with a slight body contortion. This allows us to think that the subject was still using a body referent frame.

After cardinal training, the subject made no mistakes. He used three exocentric haptic exploration patterns: "perimeter to object", "back-and-forth" and "object to object"; and only one egocentric strategy: "home base-to-object" haptic pattern.

After this cardinal points training, we observed the subject's exocentric behavior. Firstly, because he thought aloud we were able to examine a part of his spatial cognitive process based on the cardinal orientation. As we had expected, the subject only spoke about cardinal orientation. For example, he said that the triangle was in "the northwesterly corner" instead of "the top left corner". Secondly, he placed objects straight in the right position without body contortion. He seemed to make mental rotations in an easier way. Consequently this allows us to think that the subject encoded the location of the pieces in the spatial layout using an exocentric reference frame. We found these results although based on only one subject very revealing.

6. DISCUSSION

According to the previous non-visual spatial theories (Thinus-Blanc, 1997) (Hil et al, 1993) (Ungar, 2000), exocentric reference provides a higher spatial cognitive level. The cardinal concept seems to put the blind at an even more superior spatial cognitive level. As we have already seen, the compass provides available external cues (Lloyd, 2000) regardless of the subject's body position. We noticed that the subject used the tactile compass only during the first three minutes of the cardinal training but he kept answering questions about the cardinal orientations. Moreover, reaction times of the answers decreased as the training went on. This evidence supports the assumption that the subject succeeded in internalizing cardinal points in a map representation. In accordance to Vygotsky's theory, a tactile compass, as psychological instrument, reorganizes spatial cognition for our subject.

Eventually, we may take patterns on Thinus-Blanc's model of "two level spatial processing" in order to provide an explanation of spatial cognition. Our subject first used simple means to encode information in order to get acquainted with the environment. Consequently, the position of the north has been encoded from body reference. However to learn the other cardinal points specific maps are needed. The internalisation of the relationships between the north indicated by the compass and the different cardinal points proves the validity of exocentric organization as a context situated representation of the space.

7. EXPERIMENT CONCLUSION

We do not have the ambition to explain the general spatial cognition of the blind. Our point is to understand how tactile compass must be used in order to afford exocentric reference frame (*cf.* picture 4)**.** The previous theories and results lead us to think that our subject first touched the compass with egocentric haptic patterns (1) in order to encode the north direction in a body reference frame (2). After this, he associated egocentric north with exocentric cardinal map (3) stocked in long-term memory. Then he was able to use haptic exocentric strategies (4) in order to encode spatial relationships between elements (5) in a situated cardinal representation (6) on a cognitive map.

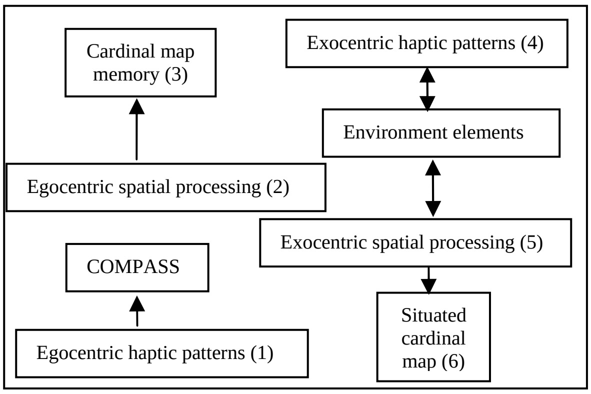


Figure 4 . *Cardinal strategy in blind condition: From egocentric pattern to exocentric haptic pattern.*

The evidence collected in this experiment supports the assumption that the use of a compass solicits exocentric patterns of exploration.

8. LIMITS AND PERSPECTIVES

We have reservations about our conclusion because of our population characteristics. Our single subject is adventitioulsy blind and is familiar with compasses, maps and sailing.On the contrary, the congenitally or adventitiously blind often have very little experience of maps and compasses. Thus our conclusion remains a hypothesis. However, we are currently conducting this experiment with twenty blind people including an experimental group and a control one. Before concluding this experiment we will emphasize that this study is a preliminary work for a more ambitious future project financed by CECIAA enterprize in CERV. In fact our aim is to understand better the spatial cognition of the blind in order to create spatial virtual reality navigation tools for them. Our experiment remains in manipulatory space, however questions about transfers between maps and largescale environment are involved. Can cardinal strategy training help blind people to improve their spatial autonomy?

How relevant is cardinal strategy for us to conceive our haptic and vocal maritime software?

9. HAPTIC AND VOCAL SIMULATOR FOR BLIND SAILORS

Usually, spatial representations can be indirectly built by symbolic media such as cartographic maps (Richardson et al, 1999). Sailboat orientation is not conceivable without maps and compasses. Even maritime spatial representation of sighted people is necessarily organized with psychological instruments.

One particular spatial feature of sailing consists in tacking when the destination is in front of the wind. When sailors zigzag, they do not follow a route but realize spatial inference tasks. In other words, if they want to reach point A, they first sail towards an imaginary point B and wait a short time and turn. This is a very a difficult situation to explain orally while the boat is heeling. Moreover, if the crew encounter rocks in their path, blind people can no longer remain at the helm. Today, accurate information is available by the means of G.P.S. (Global positioning system). However, map knowledge is required if the sailor wants to control his voyage, coordinates, bearings, distances and waypoints. In this respect, we are devoting our work to create cartographic software that will enable blind people to learn mapping and prepare trajectories.

“Most users would prefer to access tactile maps at home” (Rowell et al, 2005), that’s why we are setting up cartographic sailing simulators for blind sailors. They will be able to sail virtually with cartographic and wind constraints. Wind element and sailing principles are not complex but it is more difficult to use them in egocentric spatial representations. Our first step will be to find an easier way to teach maritime mapping - and not only maritime routing.

For a long time, sailors have employed cardinal references in order to find their way on the sea. That is why we think that our previous cardinal training task may be revised and reinvested in this project. The simplest means to test cardinal strategy and haptic exocentric patterns of exploration is to introduce a haptic device in this cartographic software. A haptic device is a “mechanical system that senses forces in remote environments and delivers those forces to the hand of the user in the form of a haptic display accessed via a rigid link” (Lederman et al, 2004). *Phantom* is a cheap available haptic device. Regarding spatial maritime layout, we will mix haptic object identification and cardinal vocal announcements. For example, blind sailors will touch a buoy and automatically hear its name. After this, if blind sailors click with the *Phantom* on another object, the announcement of cardinal orientation between these two points will be vocally announced. This is the back-and-forth haptic exocentric strategy (Thinus-Blanc, 1997) of exploration using cardinal reference. We hope blind sailors will develop new efficient strategies on this virtual sea.

We conclude maps would be better serve if used with conjunction with other multimodal devices that provide alternative sensory inputs.

Next summer, our experimentation will begin. Blind people will explore a virtual map of “the Rade of Brest” by touch. Another purpose is to find how to represent the different elements of navigation charts more intuitively. The touch of the sea will be soft and smooth, the earth will be rugged and in relief, the sailboat will be a mobile triangle, the depth will speak when you click on it, etc… Only blind sailors will tell us what works and what does not. Eventually both sighted and blind people will be able to dream together about feeling the ocean currents, the movements caused by the swell, and one day perhaps in this virtual environment we will all be able to touch a shoal of fish swimming sixty feet under the boat!

10. REFERENCES

J. Piaget & B. Inhelder. La représentation de l'espace chez l'enfant, p543, Paris, PUF, 1948.

J. M. Tellevik. Influence of spatial exploration patterns on cognitive mapping by blindfolded sighted persons. Journal of Visual Impairment & Blindness, 86, 221-224. 1992.

E.W. Hill, J. J. Rieser, M.M. Hill & J.Halpin. How persons with visual impairments explore novel spaces: Strategies of good and poor performers. Journal of Visual Impairment and Blindness. 87 (8). 1993.

C. Thinus-Blanc & F. Gaunet. Representation of space in Blind persons: Vision as a spatial sense. Journal of Visual Impairment and Blindness. 121, 1, 20-42. 1997.

J.F. Fletcher. Spatial representations in blind children. Development compared to sighted children. Journal of Visual Impairment and Blindness, 74 , 381-385. 1980.

S. Millar. Understanding and representing space: theory and evidence from studies with blind and sighted children. Oxford. Oxford University Press. 1994.

J. O'Keefe & L. Nadel. The Hippocampus as a cognitive map. Oxford. University Press. 1978.

R. Lloyd. Understanding and learning maps. *In Kitchin, R.M. and Freundschuh, S. (Eds) Cognitive Mapping: Past, Present and Future*. Routledge, London. 2000.

G. Revesz. Psychology and art of the blind. New York. Longmans, Green & Co. 1933.

P. Worchel. Space perception and orientation in the blind. Psychological monographs, 65, 332, 15. 1951.

J. Juurmaa. An analysis of the components of orientation abilityand mental manipulation of spatial relationships. Helsinski, Finland: Institute of occupational health. 1965.

R.M. Kitchin, & R.D. Jacobson, (1997). Techniques to Collect and Analyze the Cognitive Map Knowledge of Persons with Visual Impairment or Blindness : Issues of Validity. Journal of Visual Impairment and Blindness, 91 (4), 360-376.

J.J. Rieser, D.A. Guth and E.W. Hill. Sensitivity to perspective structure while walking without vision. Perception. 15(2):173-88. 1986.

R.L. Klatzky. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. *In C. Freksa, C. Habel, & K. F. Wender (Eds.), Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge* (Lecture Notes in Artificial Intelligence 1404) (pp. 1-17). Berlin. Springer-Verlag. 1998.

Y. Hatwell,. Psychologie cognitive de la cécité précoce. Paris. Dunod. 2004.

Y. Hatwell, A. Streri & E.Gentaz. Toucher pour connaître. Paris : PUF. 2000.

M.N. Avraamides, J.M. Loomis, R.L. Klatzky & R.G. Golledge. Functional equivalence of spatial representation derived from vision and language : evidence from allocentric judgements. Journal of experimental psychology : learning, memory and cognition. 30, 4, 801-814. 2004.

M. Carreiras & B. Codina. Spatial cognition of Blind and Sighted: visual and amodal hypothesis, Cahier de Psychologie Cognitive , vol. 12, 1, 51-78. 1992.

S. Ungar. Cognitive mapping without visual experience. *In Kitchin, R.M. and Freundschuh, S. (Eds) Cognitive Mapping: Past, Present and Future*. Routledge, London. 2000.

V.R. Schinazi. Spatial representation and low vision: two studies on the content, accuracy and utility of mental representations. CASA Working Papers, no.93. Working paper. Centre for Advanced Spatial Analysis (UCL), London, UK. 2005.

L.P. Acredolo, H.L. Pick & M.G. Olson. Environmental differentiation and familiarity as determinants of children memory for spatial location. Developmental Psychology, 50, 1062-1070. 1975.

R.A. Schmidt. A schema theory of discrete motor skill learning. Psychologial Review, 82, 225-260. 1975.

A.E. Richardson, D. Montello & M. Hegarty. Spatial knowledge acquisition from maps, and from navigation in real and virtual environments. Memory & Cognition, *27,* 741–750. 1999.

J. Rowell & S. Ungar. Feeling our way: tactile map user requirements – a survey. International Cartographic Conference, La Coruna, 2005.

S.J. Lederman, & R. L. Klatzky, Haptic identification of common objects: Effects of constraining the manual exploration process. Perception & Psychophysics, 66, 618-628. 2004.

ICDVRAT 2006

Session VIII. Cognitive Skills

Chair: Eva Petersson

Visual spatial search task (VISSTA): a computerized assessment and training program

A Bar-Haim Erez[1], R Kizony[2], M Shahar[3] and N Katz[1]

[1]School of Occupational Therapy, Hebrew University & Hadassah, Jerusalem

[2] Department of Occupational Therapy, University of Haifa

[3]Intelligrate, Tel Aviv, Israel

aaerez@zahav.net.il,rachelk@zahav.net.il, meir@intelligrate.biz, noomi.katz@huji.ac.il

ABSTRACT

The aim of this paper is twofold 1) to introduce a computerized platform of Visual Spatial Search Task (VISSTA), its current package and potential for a variety of additional programs, and 2) to present results of the basic package of stroke patients and healthy controls. *Method*. Participants included 39 healthy individuals; 25 patients post right hemisphere damage (RHD) with unilateral spatial neglect (USN); 27 patients post RHD without USN; and 20 patients post left hemisphere damage (LHD). All participants were tested on the computerized VISSTA and paper and pencil cancellation tests. The stroke patients were also tested on the ADL checklist and FIM. *Results*. Findings indicate that the VISSTA is a valid visual search assessment that significantly differentiated between patients following stroke and healthy controls and between different stroke patient groups. USN patients showed impairment in both visual search conditions and clear laterality bias when target was presented on the left side of a computer screen, this was true for success rate and reaction time. RHD patients without USN performed better than those with USN, however, they still show impairment in attention properties of visual search and detection of targets (both on left and right) compared to healthy individuals. *Conclusions*. The VISSTA tool was found to be sensitive to levels of visual spatial attention by means of accuracy and reaction time. Results suggest that it is important to supplement the conventional paper and pencil tests and behavioral measures with tools that provide both accuracy and RT parameters in a randomized and more complex fashion. The VISSTA is also suitable for treatment as it provides a flexible platform.

1. INTRODUCTION

Attention is defined as a cognitive mechanism that enables people to process relevant external stimuli, internal thoughts or actions while ignoring irrelevant ones (distracters). Visual spatial attention is considered the attentional mechanism that assists us in scanning the environment for relevant stimuli. Theoretical explanations of visual spatial attention mechanisms attempt to determine whether detecting a visual object occurs by simultaneously analyzing all the objects in the visual field (termed also as parallel, pop-out, or feature search), a task which requires minimal ('spread attention') or no attentional efforts, or whether visual-detection task requires successive analysis of each object (termed also as serial or conjunction search), necessitating attention shifting and selection efforts (Nakayama and Joseph, 2000; Pavlovskaya et al, 2002; Treisman and Gelade, 1980; Treismann, 1988; Treisman, 1999).

Impairments in visual search can immerge following brain damage, either due to impaired spatial attention mechanisms like in the complex neurological disorders of Unilateral Spatial Neglect (USN), or field cut deficits such as hemianopsia. Visual spatial deficits affect the ability to function in many aspects of daily living and have serious consequences for rehabilitation and long term functional capacity (Heilman et al, 1993; Kalra et al, 1997; Katz et al, 1999), for this reason it is essential to assess and treat these disorders in the rehabilitation process.

Assessment of visual spatial deficits (including USN) is usually performed using paper and pencil tests. Various cancellation tasks are employed where patients are asked to search for a specific target amongst varied number of distracters (Weintraub & Mesulam, 1987). The conventional tests have some limitations, usually the tasks do not change from one evaluation to another and have only one level of difficultly that may lead to a ceiling effect, in addition, reaction time, which is an important component of visual spatial attention, is not provided. Moreover there is a need to provide clinicians with user friendly training tools in order to improve their patients' visual search abilities in a controlled, gradable way. There are conflicting results with regards to the performance of patients post stroke in the two levels of visual search (feature versus conjunction). Some studies show that patients with USN have difficulties detecting targets in the contralesional side with significantly slower reaction time (RT) at both levels, while some claim that the main difficulty is in the more complex serial search (Behrmann et al, 2004).

The VISSTA, computerized platform, was developed to serve both assessment and training of visual spatial attention deficits. VISSTA is based on Sun Micro System's Java platform. It requires initial installation of Java's virtual machine (which is freely available). After this one time installation, it can be easily used by launching the application. VISSTA manages a database of the patients and the sessions. This detailed database enables retrieval of information and report generation. The reports can be generated according to the user's preferences (level of details, output format etc.) and either be printed or saved. The system also enables access to the raw data for further analysis. At the heart of VISSTA's operation stands the loop of presenting a stimuli (a screen) and accepting the patients input (while measuring its time and correctness).

To facilitate a variety of different type of stimuli VISSTA has a powerful feature called packages. Packages are easy to distribute files that contain the full media (images) and definitions (grouping, desired response, durations etc.) that are required to run a session. VISSTA can easily edit, create and read packages, hence supporting a quick, simple and easy enhancement of the available content. The base package was developed based on Treisman and Gelade feature theory, which is comprised of two main conditions; feature and conjunction search, 108 screens are randomly arranged with variations of location of target, number of distracters, side of the screen etc. and approximately 20% of the screens are without target (only distracters).

The newly developed computerized visual search test and training program VISSTA (Visual Spatial Search Task), presented in this paper, applies both 'feature' and 'conjunction' search principles to assess identification rate and reaction time (RT) in the two conditions. The aims of this paper are to report the psychometric properties of the VISSTA and to discuss the advantages of using computerized search tasks, both feature and conjunction based, in the assessment of USN.

2. METHODS

2.1 Participants

Healthy participants (N=100), ages 25-90 years old, independent in activities of daily living (ADL), were recruited from the community and tested for standards on VISSTA. Thirty nine of them were selected from the pool to match the patients' age group (mean age = 63.5, SD=13.6). Participants post stroke were recruited from Lowenstein rehabilitation center and than allocated to three groups: a) 25 patients with right hemisphere damage (RHD) with unilateral spatial neglect (RHD +USN); b) 27 patients with right hemisphere damage (RHD) without unilateral spatial neglect (RHD – USN); and c) 20 patients with left hemisphere damage (LHD) without USN. Patients with visual field deficit were excluded from the study. All patients were 3-12 weeks post their stroke event.

2.2 Instruments

2.2.1 Measures to determine USN. Within the RHD patients the diagnosis of USN was determined based on deficits in the following measures: Behavioral Inattention Test (BIT) (Wilson et al, 1987), a widely used battery of tests for assessing spatial deficits and screening for USN; The Mesulam and Weintraub random symbol cancellation task (MWCT) (Weintraub, 2000); Catherine Bergego Scale (ADL checklist) (Azouvi, 2003) a report of an experienced occupational therapist concerning USN behavior during basic ADL activities; and the Functional Independence Measure-FIM (Granger et al, 1993) is used to measure the degree of disability and burden of care in everyday activities.

2.2.2 The computerized visual search task (VISSTA). The program includes two conditions, feature and conjunction. Feature visual search: subjects were asked to detect a visual target (red circle) located amongst several peripheral distracters on a computer screen. The target differed from the distracters by color (e.g., a red circle among blue circles). The target appeared randomly in one of 20 predetermined locations (5 targets on each quarters of the screen: right / left; up / down) and 5 screens without targets. The distracters appeared randomly on both sides of the screen. The location of target and the set-size of distracters varied (from 3 distracters to 23). Thirty percent of the screens were without target (presenting only distracters and termed 'catch trials'), overall there are 108 screens that the participants has to respond to. The subjects were required to press one button as soon as the target was detected and another button if no target was detected. Each screen appeared for 3000 msec regardless of participant's response. Reaction time and success rates were measured. Conjunction visual search: Subjects were asked to detect the same visual target (red circle), however this time, among two types of distracters (blue circles and red squares), the target differed from the distracters by at least one primary feature (e.g., color or shape) but was similar on the other feature. The location of target and the distracters protocol were the same as for the 'feature' condition. The exposure time was 4500 msec based on a preliminary trial made with healthy elderly and patients post stroke. See Figure 1 for an illustration of one screen.

For training purposes the parameters of the program can be changed and graded, namely the exposure time, the number of screens, the stimuli etc. The output provides detailed data on rates of success and omissions, divided by quadrants and response time.

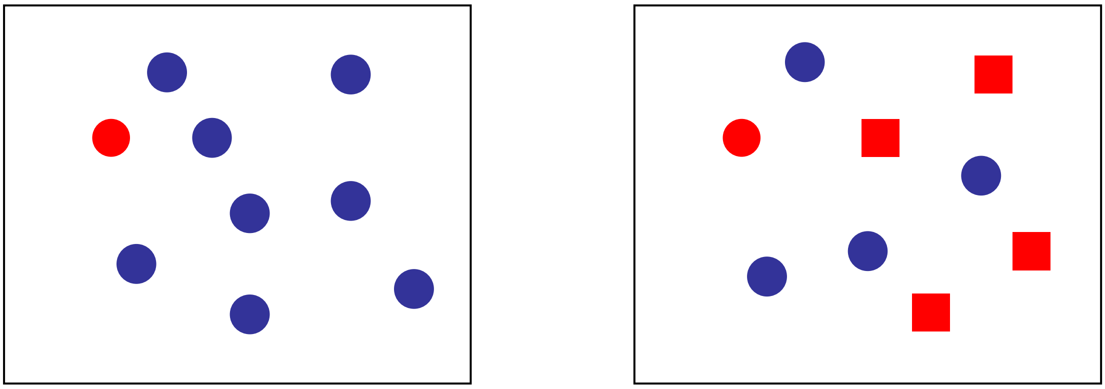


Figure 1: *Feature and conjunction search.*

2.3 Procedure

Patients first went through the paper-and-pencil tests to determine the presence of USN. Patients were allocated to one of three groups according to hemisphere lesion side and the evidence of USN. The participants were seated in front of the monitor (size 17inch) at their arm length distance. All the participants (healthy and patients) went through training on the computerized tests to make sure they understood the instructions and were capable of performing the two tasks. The computerized visual search assessment was conducted in one session with a short break between the two tasks. The feature condition was preformed first and than the more complex conjunction condition. For the patients the response buttons were placed in front of the subjects in a position that was comfortable and enabled them to press the two buttons without looking down. Healthy subjects used the keyboard. The FIM and ADL checklist were completed by the treating therapists. The research study received the Institutional Review Board (IRB), Human Rights Committee approval and all participants signed a consent form before entering the study.

3. RESULTS

3.1 Visual search computerized tasks: comparison between and within groups.

The data was analyzed using One Way ANOVA and coefficient contrasts tests to compare the groups in respect to each condition, feature and conjunction. The analysis was conducted separately for the responses made to targets on left and right visual fields and where no target was presented (catch trials). All ANOVA results between groups were significant at p=.026-.0001 (see Tables 1 and 2). Results for each condition with coefficient contrasts show that RHD +USN patients preformed significantly worse than the other groups in all conditions except for the feature condition when targets were presented in the right visual field.

3.2 Feature versus conjunction.

In all four groups RT was longer for the 'conjunction' compared to the 'feature' condition (RHD+USN: z = 3.7, p = .007; RHD-USN: z = 4.5, p = .000; LHD: z = 3.8, p= .001; Healthy: z = 5.4, p = .000; (see Table 2). In RHD groups, with and without USN, patients' success rate was higher for 'feature' condition than

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

‘conjunction’ (RHD+USN: z = 3.1, p = .002; RHD-USN: z = 2.4, p = .015). In the LHD group a significant difference was detected in the conjunction condition such that the RT for right targets was longer than RT to left targets (z = 2.3, p = .02). Both comparisons between and within groups are demonstrated in Figures 2 and 3.

Table 1: *Hit rate (%) in feature and conjunction search; One way ANOVA between groups.*

Group		**Feature**	
	Target Left	Target Right	Catch Trials
Healthy	97.9 (.03)	95.9 (.06)	97.5 (.06)
LHD	90.2 (.21)	86.9 (.19)	86.4 (.22)
RHD-USN	83.2 (.24)	88.6 (.15)	82.8 (.24)
RHD+USN	52.9 (.33)	85.8 (.15)	72.0 (.23)
F (p)	23.2 (.000)	3.7 (.014)	9.7 (.000)
		Conjunction	
	Target Left	Target Right	Catch Trials
Healthy	92.6 (.1)	92.3 (.11)	92.2 (.12)
LHD	93.7 (.07)	94.4 (.06)	92.8 (.09)
RHD-USN	85.5 (.13)	88.9 (.13)	80.4 (.22)
RHD+USN	40.9 (.27)	76.8 (.18)	67 (.19)
F (p)	61.7 (.0001)	9.6 (.0001)	14.3 (.0001)

Table 2: *Reaction time (msec) in feature and conjunction search; One way ANOVA between groups.*

Group		**Feature**	
	Target Left	Target Right	Catch Trials
Healthy	940.3 (451.4)	983.7 (419.7)	1122.8 (420.8)
LHD	1090.5 (295)	1188.7 (282.9)	1515.6 (580.4)
RHD-USN	1378.3 (977.4)	1224.9 (444.5)	1478.8 (419.8)
RHD+USN	2522.8 (1261)	1421.1 (912.6)	1806.3 (682.1)
F (p)	20.5 (.0001)	3.2 (.026)	9.2 (.0001)
		Conjunction	
	Target Left	Target Right	Catch Trials
Healthy	1455.1 (616.7)	1589.8 (645.9)	2028 (506.9)
LHD	1697.3 (447.5)	1919.2 (522.9)	2643.6 (483.2)
RHD-USN	2024 (706.6)	1910.2 (699.5)	2537.8 (697.9)
RHD+USN	3701.9 (1000.1)	2255.6 (906.3)	2935.8 (1223)
F (p)	53.5 (.0001)	4.6 (.005)	7.8 (.0001)

RHD/LHD = Right- / Left-hemisphere damage; USN+/- = Patients with/without unilateral spatial neglect. Catch trials = Trials where no target was presented.

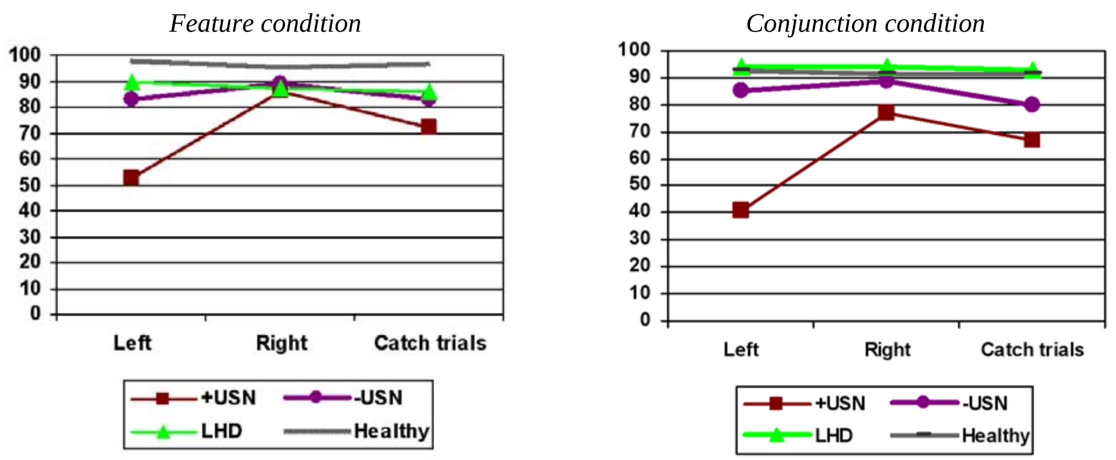


Figure 2: *Hit Rate (percent success) in Feature and Conjunction.*

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

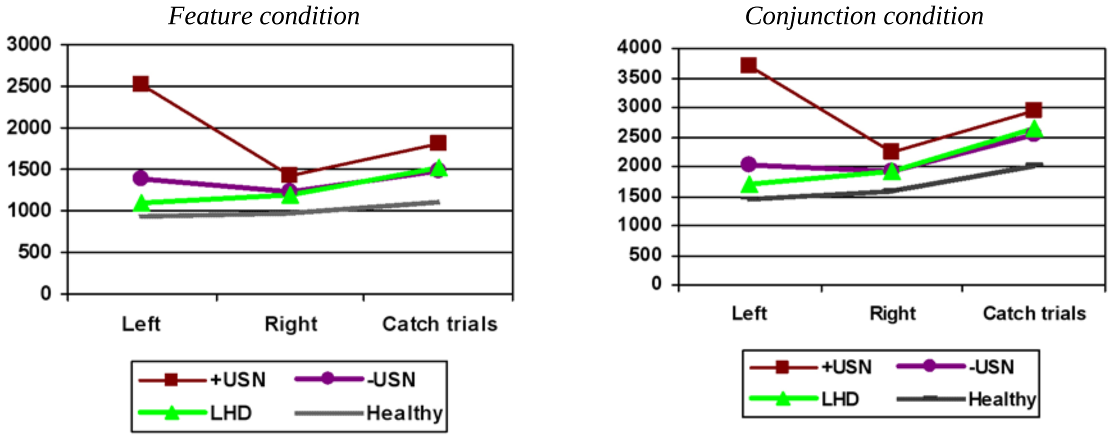


Figure 3: *Reaction Time in Feature and Conjunction.*

3.3. *Correlations between the visual search conditions and USN measures*

Spearman rho correlations between VISSTA and USN paper and pencil measures show moderate significant correlations at $p < .001$ both in the 'feature' and 'conjunction' conditions. In the RHD+USN (feature $r = .62$; conjunction $r = .57$), healthy (feature- $r = .49$; conjunction- $r = .46$) and LHD (feature- $r = -.56$; conjunction $r = .66$) groups, no significant correlations were found in the RHD-USN. Correlations with the ADL checklist (that was relevant only to the USN patients) revealed the same trend for the RHD+USN group (feature- $r = -.71$, $p<.001$; conjunction $r = -.51$, $p < .005$) indicating that higher score of USN in activities of daily living (meaning more disability) correlated with lower success rate in visual search. Correlations between the FIM measure and the total success rate in each visual search condition show significant moderate correlations in the RHD+USN group for both conditions (feature- $r = .46$, $p < .005$; conjunction $r = .65$, $p < .001$).

4. CONCLUSIONS

The computerized VISSTA program is based on a visual search theory differentiating between levels of visual attention – simple feature search versus a more complex attentional driven conjunction search. Findings indicate that the VISSTA is a valid visual search assessment that significantly differentiated between patients following stroke and healthy controls and within stroke patients groups. VISSTA scores were moderately correlated with the paper and pencil and ADL tests in both feature and conjunction conditions. The computerized version provides additional information on visual search abilities and attention by measuring the reaction time in addition to successful detection of stimuli. The VISSTA program provides additional important parameters for assessing and treating visual spatial attention such as limiting the exposure time of the stimuli presentation as well as providing repeated presentation of randomized stimuli over relative long periods of time (5.5-8 minutes) which challenges sustained attention capabilities.

The conventional paper and pencil tests have some limitations compared to a computerized version. Usually the tasks in the paper and pencil tests do not change from one evaluation to another and have only one level of difficultly that may lead to a ceiling effect, while in a dynamic platform, such as a computerized tool, it is possible to change the stimulus, its background and its exposure time thus increasing complexity and sensitivity. In addition, the measurement of reaction time provides important information on spatial attention which is not provided by the conventional tests (Deouell, 2005). The computerized test seems to provide valid and sensitive information that can be useful for the therapist to assess severity of visual spatial deficits and detect even subtle changes in visual attention prior to further evaluation.

In conclusion, the VISSTA tool was found to be sensitive to levels of visual spatial attention by means of accuracy and reaction time. Results suggest that it is important to supplement the conventional paper and pencil tests and behavioral measures with tools that provide both accuracy and RT parameters in a randomized and more complex fashion. The computerized visual search tool can serve as a sensitive measure for the presence of visual spatial deficits (and general decreased attention) at various levels in the rehabilitation process in addition to paper and pencil and functional tests.

Acknowledgement: The initial development of the VISSTA and this study was funded by the scientific brunch of the Israeli Ministry of Health.

5. REFERENCES

K Nakayama and J S Joseph (2000), Attention, pattern recognition and Pop-Out in visual search. In: R Parasuraman ed. *The Attentive Brain*, Bradford Books: Cambridge pp 279-298.

M Pavlovskaya, H Ring, Z Groswasser and S Hochstein (2002), Searching with unilateral neglect. *J. cog. Neusosci.,* 14 pp. 745-756.

A Treisman and G Gelade (1980), A feature integration theory of attention. *Cognit. Psychol*. 12, pp. 97-136.

A Treismann (1988), Feature and objects: the 14th Barlett memorial lecture. *Quarterly J. Exp. Psych*. 40, pp. 201-237

A Treisman (1999), Feature binding, attention and object perception. In: G W Humphreys J Duncan and A Treisman eds. *Attention, Space and Action.* Oxford University Press, pp. 91-111.

K M Heilman, R T Watson and E Valenstein (1993), Neglect and related disorders. In: K M Heilman and E Valenstein eds. *Clinical Neuropsychology* (Third edition). Oxford University Press: New York, pp. 279-336.

L Kalra, I Perez, S Gupta and M Wittnik (1997), The influence of visual neglect on stroke rehabilitation. *Stroke*, 28, pp. 1386-1391.

N Katz, A Hartman-Maeir, H Ring and N Soroker (1999), Functional disability and rehabilitation outcome in right hemisphere damaged patients with and without unilateral spatial neglect.*Arch. Phys. Med. Rehabil.,* 80, pp. 379-384.

S Weintraub and M M Mesulam (1987), Right cerebral dominance in spatial attention: Further evidence based on ipsilateral neglect. *Arch. Neurol*., 44, pp. 621-625.

M Behrmann, P Ebert and S E Black (2004), Hemispatial neglect and visual search: A large scale analysis. *Cortex*, 40, pp. 247-263

S Weintraub (2000), Neuropsychological assessment of mental state. In *Principles of Behavioral and Cognitive Neurology* (Second edition), M M Mesulam Ed. Oxford University Pres: Oxford, pp. 121-173.

B A Wilson, J Cockburn and P W Halligan (1987), The development of a behavioral test of visuospatial neglect. *Arch. Phys. Med. Rehab.* 68, pp. 98-102.

P Azouvi, S Olivier, G de Montety, C Samuel, A Louise-Dreyfus and L Tesio (2003), Behavioural assessment of unilateral neglect: Study of the psychometric properties of the Catherine Bergego Scale. *Arch. Phys. Med. Rehab*., 84, pp. 51-57.

C V Granger, A C Cotter, B B Hamilton, and R C Fiedler (1993), Functional assessment scales: A study of persons after stroke. *Arch. Phys. Med. Rehab*., 74, pp. 133-138.

L Y Deouell, Y Sacher and N Soroker (2005), Assessment of spatial attention after brain damage with dynamic reaction time test. *J Int Neuropsychol Soc*., 11, pp.697-707.

Analysing the navigation of mentally impaired children in virtual environments

C Sik-Lányi[1], R Mátrai[2] and I Tarjányi[3]

Department of Image Processing and Neurocomputing, University of Pannonia,
Egyetem Street 10, Veszprém, HUNGARY

[1]*lanyi@almos.vein.hu*, [2]*ritka@vision.vein.hu*, [3]*istvan_tarjanyi@vipmail.hu*

www.vein.hu, www.knt.vein.hu, vision.vein.hu

ABSTRACT

In today's information society, computer users frequently need to seek for information on home pages as well as to select among software functions. A well-designed interface is essential in order to find everything necessary and meet the requirements of both the average user and users with special needs. Our project set out to discover where and with how much contrast objects should be placed on the screen in order to find everything easily. We examine what kind of characteristic searching routes can be found and whether we can find differences between the average user and mentally retarded user in navigation and everyday searching exercises.

1. INTRODUCTION

Numerous visual experiments have been carried out to discover how to organize objects and how to set up colours to help the process of searching. In the case of visual searching exercises, elements like orientation, colour, motion, size, shape, position of objects and of the background, density of objects (spatial frequency) could facilitate or hinder the search. During the past 25 years numerous visual search models have been produced, which dealt with the question of how humans are able to find objects, and tried to predict the results of visual search tasks. (Chen 1982, Mátrai et. al. 2005, Pomerantz and Pristach 1989, Sagi 1988, Treisman and Gelade 1980, Treisman and Gormican 1988, Wolfe 1994)

Visual search models distinguish two stages in searching. The first is a preattentive, largely parallel stage that processes information about basic visual features (e.g. colour, form, depth) across large portions of the visual field. Information processing results in the formation of feature maps in the brain. A subsequent limited-capacity stage follows in which other, more complex operations take place (e.g. face recognition, reading) over a limited portion of the visual field. (Treisman and Gelade 1980, Wolfe 1994).

An interesting question is how humans search for information on homepages among a lot of text, menus and pictures. Studies by Nielsen (2006) showed that users do not read but scan homepages: they scan the page and try to select information from it which can be useful for them. Nielsen's recent studies on eye-tracking discovered patterns in the way pages were scanned. Three types of web pages were examined. The first one was an "About Us"-page and was scanned in an L shape which flipped horizontally. On a product page on an e-commerce site, the scanning pattern resembled an F. Search engine results were scanned in an E shape.

In our current research we investigated whether similar shapes can be found for average and mentally retarded children in playful search tasks. This research was more complex than Nielsen's tests, because the found objects had to be clicked and the order of the perceived and clicked objects is not necessarily the same.

Nowadays interactive (dynamic) homepages (e.g. forms, web-shopping pages, on-line ticket ordering pages etc) are frequently used. In these one has not just to read, but has to click on some objects, in the correct order. We wanted to discover the sequence of the clicking in the various tasks. 15 mentally impaired and 50 average children took part in the research.

2. METHOD

In order to examine the visual search process several programs were developed, from which the tasks introduced in this article were also realized in a 3D environment. Effects of each visual feature on reaction time and searching routes can be described with a linear regression model in which the input variables are the visual properties and the response variable is the latency to click. Results will be evaluated with the help of this regression model. The following hypothesis will be tested:

- H1 hypothesis: Properties of objects (direction, colour, placement, etc.) have different effects on reaction time and searching route of mentally impaired users.
- H2 hypothesis: In 3D space reaction time is longer for the same search task and the visual properties of objects have different effects on reaction time and searching route.

2.1 Presentation of the software

In the first exercise the same object is shown four times on the screen in a row each time in one of four different orientations, facing to the left, to the right, up or down (Figure 1). In one worksheet only one figure appears. Orientation is generated randomly at the beginning of each trial. In the first part of the exercise figures facing to the left should be painted red while in the second part the ones facing to the right should be painted blue. On some of the trials none of the figures looks to the left and/or to the right.

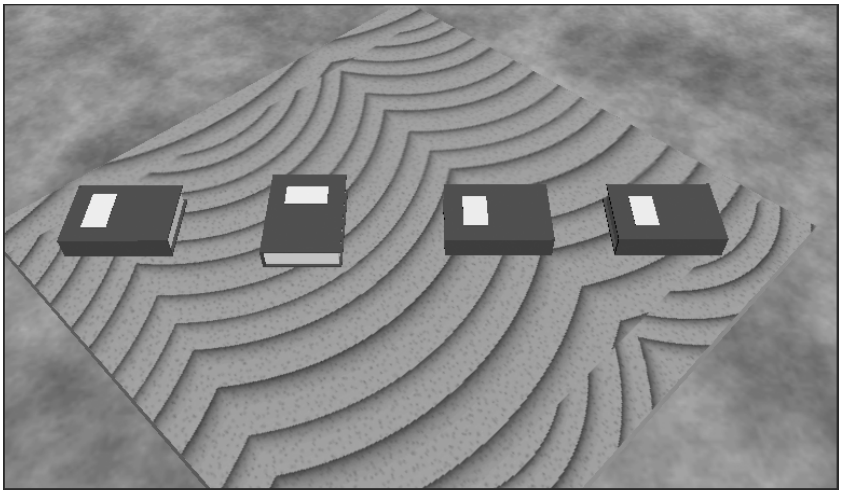


Figure 1. *A worksheet from the 3D-version where figures looking left should be found.*

Previous experiments found that if the object to be found is not present, searching time increases (Wolfe 1994). We expect similar results in the current test. As not only average children are involved in these examinations but also mentally retarded children, it was important to prepare simple exercises. We also compared the searching time for figures facing left and right in case there were more of one than the other; the order of clicking on them; and, in case the target figure was not displayed, the amount of time taken to realise this.

In the following task shapes (star, circle, square, etc.) were cut from a rectangle, and put next to the rectangle with some other similar shapes. The task was to put back the shapes. We examined whether position and form influenced the order in which shapes were chosen.

Several exercises were constructed to examine the visual search process when the target is the type of shape and the displayed objects are geometrical figures. Each worksheet has also been prepared in 3D form. We examined how results of searching depend on the contrast between foreground and background, and how they differ between special needs users and average users. We also examined whether some kind of order can be found in search routes or whether they are completely random, and also compared 3D and 2D tasks.

2.2 Participants

The programs have been tested in two primary schools where during computer aided and conductive lessons pupils could play with the program with the guidance of their teachers. One school was for mentally retarded children while the other was for average pupils.

2.3 Devices

Seventeen inch CRT-monitors were used in the school computer rooms. Teachers took care that the children did not bend too close to the monitor in order to maintain a focal distance of 60 cm. As understanding and remembering directions is often problematic for mentally impaired children we used a trick: on the top edges of each monitor the word LEFT or RIGHT indicates the directions accordingly. This helped them when completing the exercises involving differentiation of orientation.

3. RESULTS

In this research we intended to find the most salient properties of the object (placement, direction, etc.), which influenced the reaction time both for mentally impaired and average children. Since 3D displays are becoming more and more popular it was important to compare the reaction time of similar tasks in 2D and 3D. However there was additional information in the 3D tasks (spatial placement, highlights and shadows), which facilitated or hindered the navigation of the task, therefore it was necessary to compare the results of tasks in 2D and 3D.

3.1 The effect of object placement on navigation

In the first visual search experiment the relationship between placement of objects and order of clicks was examined. The first object was at the extreme left side and the fourth object was at the extreme right side of the display (see figure 1). The effect of reading direction had a significant effect on clicking. Generally the first clicked object was the first object, in other words the first object reached if scanning from left to right. The second clicked was the second object and so an. However, for the mentally impaired children the effect of the orientation was more significant than the position of the object. Since the orientation of the object is a random variable, results are presented separately for each orientation (left, right or up) of the object. Figure 2 shows the relative frequency of the first clicked object. Here the right answers are the first boxes, because the task was to click the objects, which were looking left.

In Figure 2 the first column on the left shows that the first click occurred on the first figure (from the left) in 68% of all tasks where the first figure faced towards the left. The second column shows that the first click occurred on the first figure in 30% of all tasks where the first figure faced right, etc. Since the task was to find figures facing to the left the correct answers are indicated by the first boxes. In this comparison it can be seen that the number of mistakes was lower for average children, and the effect of position was more significant than orientation of the object for mentally impaired children. The figure on the extreme left was clicked first even if it was in the wrong orientation.

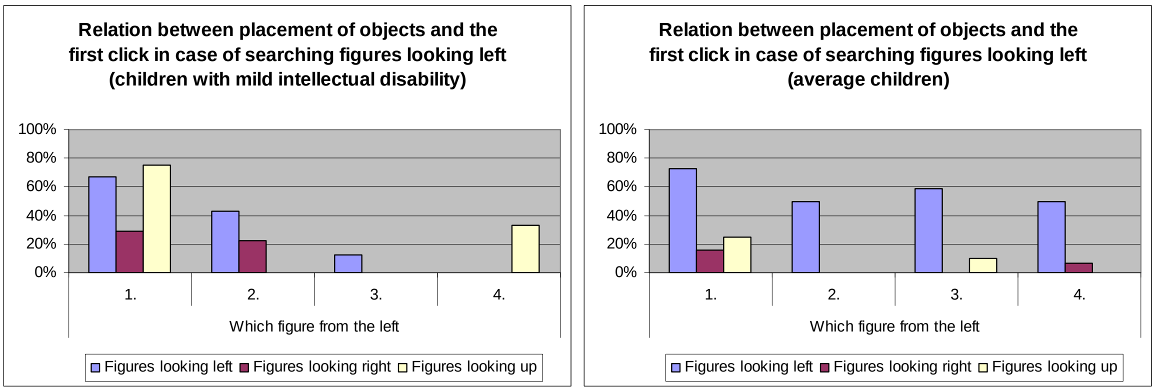


Figure 2. *First clicked objects where the task was to find figures facing left.*

In the second half of the game the task was to find objects which were facing to the right.

The effect of direction of reading is very significant for mentally impaired children. They made fewer errors (the right answers are indicated by the second column), but in the mentally impaired children the effect of solving the previous task (where they had to find objects facing left) has influenced the results. There are some cases where children clicked the objects facing left (Figure 3).

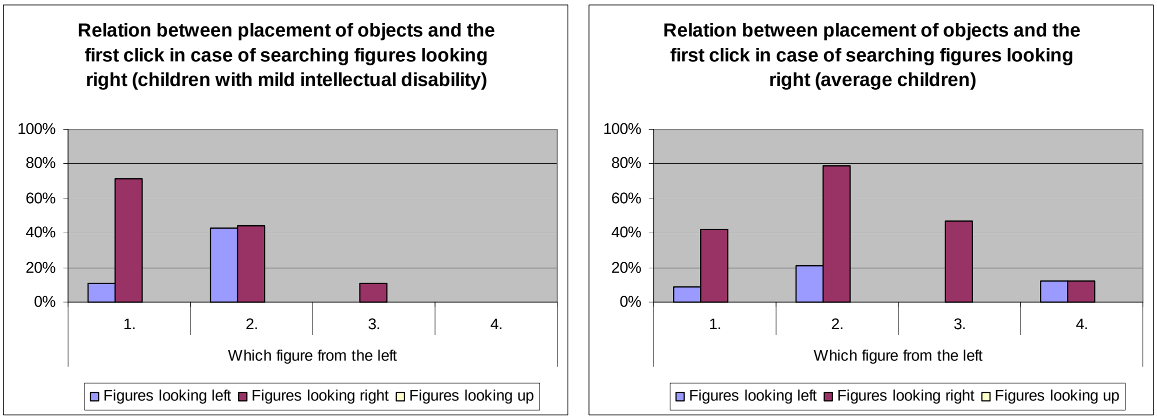


Figure 3. *First clicked objects where the task was to find figures facing right.*

3.2 Finding the right place for the shapes

In this task the order of selecting objects was examined. Figure 4 shows which figure was selected first.

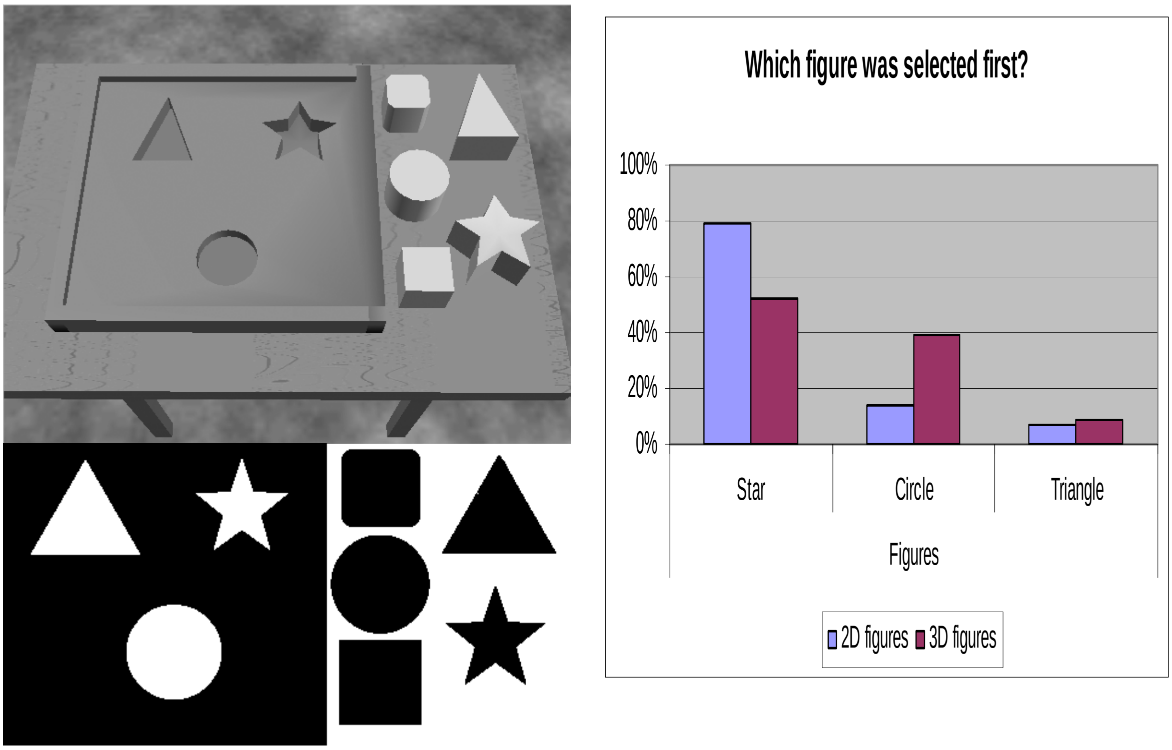


Figure 4. *Which figure was selected first?*

It is understandable that the first selected object was the star, because this object is very different from the other objects. However, it is very interesting that the triangle was selected by very few children as the first object as it is the next most different object after the star. The reason for this phenomenon might be the position of the objects: the triangle was on the right hand side, the circle on the left hand side, and children probably started from the left.

3.3 Finding geometrical shapes

This task was more difficult than those discussed previously. In this 2D and 3D worksheet children had to find geometrical shapes. The aims of this investigation were to find typical search routes, and to compare the reaction times and search routes in 2D and 3D tasks.

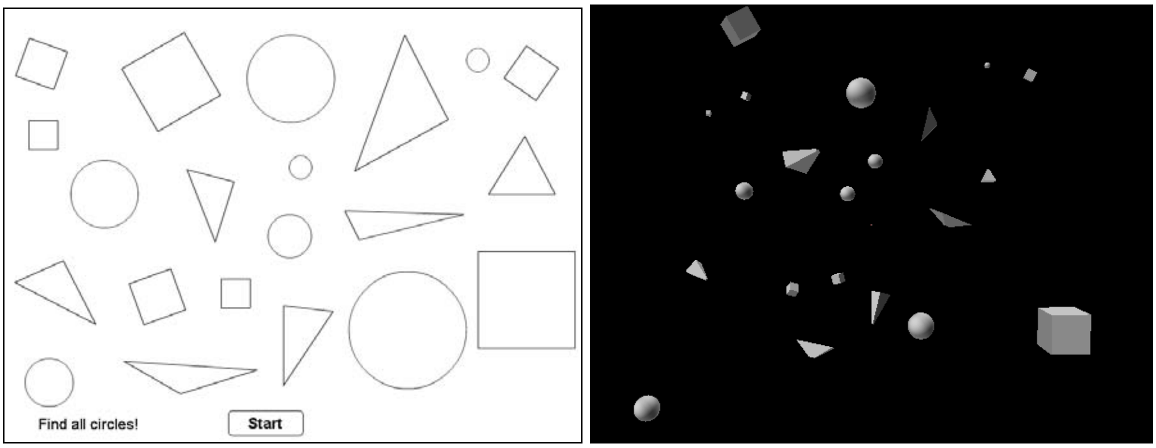


Figure 5. *Two similar tasks in 2D and 3D where circles/spheres had to be found.*

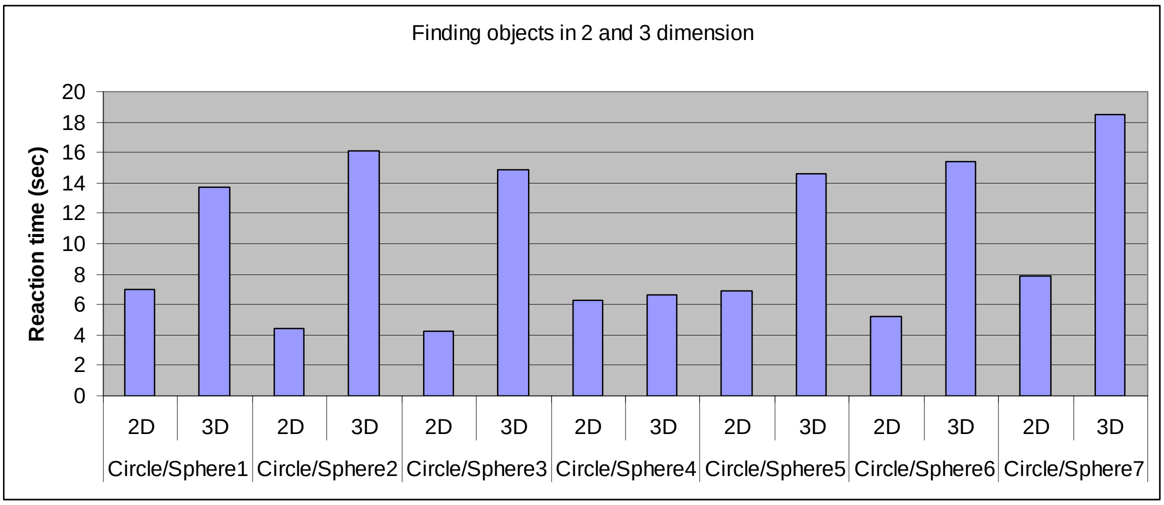


Figure 6. *Comparison of reaction times for the similar tasks in 2D and 3D.*

Figure 6 shows that the reaction times for the 3D tasks are higher than those for the 2D tasks. This may be because more visual features had to be processed in the three dimensional task.

Figure 7 shows two typical search routes in both tasks.

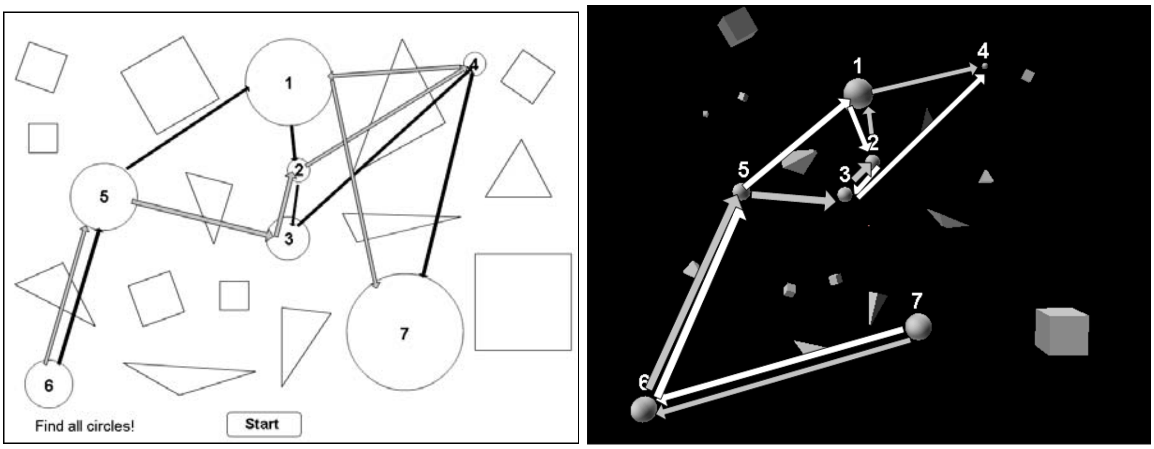


Figure 7. *Comparison of searching routes of the similar tasks in 2D and 3D. (Different colours of arrows show the alternative searching routes).*

The “left to right” direction of the search can be seen in all search routes. Since the *Start* button is at the bottom of the worksheet the direction of search is upwards. It is interesting that in the 2D task finding the 7th circle was the most difficult. This might be because it was near the bottom of the right hand side where, according to Nielson, people rarely look. It might be that after focusing on the small circles it could be more difficult to observe it because of its large size.

4. CONCLUSIONS

The position of figures is very important in navigation and finding objects. Generally this effect is more significant than the shape or direction of figures. The direction of navigation is usually the same as the reading direction, but this effect is more significant in mentally impaired children. In 3D navigation there are additional effects (lights, shadows, etc.), which usually make it more difficult to find objects.

Acknowledgements: The authors would like to say thank to two teachers, Lilla Vági (Kozmutza Flóra Primary School, Veszprém) and Endre Buday (Teleki Blanka Primary and Secondary Grammar School, Székesfehérvár) for helping the organization of the experiments.

5. REFERENCES

L Chen (1982), Topological structure in visual perception, *Science*, **218**, pp. 699-700.

R Mátrai, C Sik-Lányi, Zs T Kosztyán (2005), Games Test to Study the Navigation of Handicapped Children, In proceedings of *CEMVRC (Central European Multimedia and Virtual Reality Conference) 2005, 8-10 June 2005, Prague, Czech Republic*, pp. 167-168.

J Nielsen (2006), F-Shaped Pattern For Reading Web Content, *Jakob Nielsen's Alertbox, April 17, 2006*, http://www.useit.com/alertbox/reading_pattern.html

J R Pomerantz and E A Pristach (1989), Emergent features, attention, and perceptual glue in visual form perception, *Journal of Experimental Psychology: Human Perception & Performance*, **15**, pp. 635-649.

D. Sagi (1988), The combination of spatial frequency and orientation is effortlessly perceived, *Perception & Psychophysics*, **43**, pp. 601-603.

A Treisman and G Gelade (1980), A feature integration theory of attention, *Cognitive Psychology*, **12**, pp. 97-136.

A Treisman and S Gormican (1988), Feature analysis in early vision: Evidence from search asymmetries, *Psychological Review*, **95**, pp. 15-48.

J Wolfe (1994), Guided Search 2.0. A revised model of visual search, *Psychonomic Bulletin & Review*, **1** (2), pp. 202-238.

Tangible user interfaces: tools to examine, assess and treat dynamic constructional processes in children with developmental coordination disorders

S Jacoby[1,2], N Josman[1,3], D Jacoby[1,4], M Koike[5,6], Y Itoh[5,7], N Kawai[5,8], Y Kitamura[5,9], E Sharlin[10] and P L Weiss[1,11]

[1]Department of Occupational Therapy, University of Haifa, Haifa, ISRAEL

[5]Human Interface Engineering Laboratory, Graduate School of Information Science and Technology, Osaka University, Osaka, JAPAN

[10]Interactions Laboratory, Department of Computer Science, University of Calgary, Calgary, Alberta, CANADA

[2]*sigalj1@012.net.il*, [3]*naomij@research.haifa.ac.il*, [4]*dojacoby@cisco.com*, [6]*koike.minori@ist.osaka-u.ac.jp*, [7]*itoh@ist.osaka-u.ac.jp*, [8]*kawai.norifumi@ist.osaka-u.ac.jp* [9]*kitamura@ist.osaka-u.ac.jp*, [10]*ehud@cpsc.ucalgary.ca*, [11]*tamar@research.haifa.ac.il*

[1]*hw.haifa.ac.il/occupa/LIRT*, [5]*www-human.ist.osaka-u.ac.jp/*
[10]*pages.cpsc.ucalgary.ca/~ehud/*

ABSTRACT

Tangible User Interfaces (TUIs) are a subset of human-computer interfaces that try to capture more of the users' innate ability of handling physical objects in the real world. The TUI known as ActiveCube is a set of graspable plastic cubes which allow the user to physically attach or detach cubes by connecting or disconnecting their faces. Each cube is essentially a small computer which powers up and communicates with its neighbours upon connection to a neighbouring cube. When users assemble a physical shape using the system they also connect a network topology which allows ActiveCube to digitize and track the exact 3D geometry of the physical structure formed. From the user's perspective, ActiveCube is a very powerful tool; the 3D shape being built with it physically is tracked in the virtual domain in real-time. ActiveCube's use as a concrete, ecologically valid tool to understand dynamic functional processes underlying constructional ability in either typically developed children or in children with neurological pathology has not yet been explored. The objective of this paper is to describe the ActiveCube interface designed for assessing and treating children with Developmental Coordination Disorder (DCD). In our pilot study, six male children, aged 6 to 7 years, three with DCD and three who are typically developed were tested. The children's task was to successively use the ActiveCubes to construct 3D structures in a "matching" strategy known as "Perspective Matching". The usability results showed that all the participating children enjoyed the tasks, were motivated and maintained a high level of alertness while using the ActiveCubes. More than 80% of them found the tasks to be easy or moderate. "Similarity" data from single subjects has been used to show differences in constructional ability between children with DCD and those who are typically developed. This automated ActiveCube three-dimensional (3D) constructional paradigm has promise for the assessment and treatment of children with DCD.

1. INTRODUCTION

Tangible User Interfaces (TUIs) are a subset of human-computer interfaces that try to capture more of the users' innate ability of handling physical objects in the real world. Ullmer and Ishii (2001) define TUIs as "devices that give physical form to digital information, employing physical artefacts as representations and controls of the computational data". We highlight a subset of TUIs which we call spatial TUIs: interfaces that mediate interaction with shape, space and structure. We believe efficient spatial TUIs offer intuitive spatial mapping between their physical and digital qualities, unify input and output space and enable intuitive

 ISBN 07 049 98 65 3

trail-and-error activity (Shalin et al, 2004). The highly interactive and spatial nature of TUIs, as well as their current relatively large size, motivated several research efforts that mapped TUI applications to children games and playing activity. Triangles are flat TUIs which allow interactive construction of 2D shapes. Triangles were used for creating interactive narrative of a nonlinear story, allowing the users to control the story's progress as well as parts of its content by physically manipulating, connecting and disconnecting the Triangles (Gorbet et al, 1998).

An example of a spatial TUI that supports 3-dimensional (3D) construction is the ActiveCube system (Kitamura et al, 2001), shown in Figure 1. ActiveCube is a set of graspable plastic cubes (the current version dimensions are 5 cm per edge) which allow the user to physically attach or detach cubes by connecting or disconnecting their faces. Each cube is essentially a small computer which powers up and communicates with its neighbours upon connection to an adjacent cube. When users assemble a physical shape using the system they also connect a network topology which allows ActiveCube to digitize and track online the exact 3D geometry of the physical structure formed. From the user's perspective, ActiveCube is a very powerful tool; the 3D shape being built with it physically is tracked in the virtual domain in real-time.

Figure 1. *The ActiveCube system showing connections to run with a laptop computer.*

ActiveCube supports a rich variety of onboard input and output devices. For example, cubes can be equipped with a gyroscopic sensor which tracks the structure's 3D orientation. Cubes can also be equipped with a light source that can be switched to illuminate them, as well as with touch sensors, vibrators, a motorized propeller and a variety of other sensors and actuators. The current ActiveCube implementation links to a host PC through a tethered cube, or base cube, which controls the network and is the first cube of each apparatus constructed with the system (Watanabe et al, 2004).

The ActiveCube TUIs have been used as the basis for the development of Cognitive Cubes, a system that investigates adult spatial cognitive ability (Sharlin, et al., 2002). The system projected a 3D virtual prototype on a large screen and the participant was asked to construct the virtual prototype using the ActiveCube physical blocks. As the participant progressed the system extracted in real-time the geometry of the structure and analyzed the similarity between the physical construction and the virtual prototype. Testing showed that Cognitive Cubes were sensitive to age as well as to dementia (Sharlin et al, 2002).

ActiveCube has also been used to explore how human subjects use this novel computer interface to interact with narrative software. TSU.MI.KI is a novel toy based on the ActiveCube TUI (Itoh et al, 2004), practically a technological enhancement of the Japanese classic tsumiki (a traditional set of wooden blocks). TSU.MI.KI allows children to actively move through a story space by physically interacting with a set of physical cubes. The TSU.MI.KI player is confronted with several tasks through a story narrative displayed on a computer screen. In order to confront the story puzzles and challenges the child needs to play with their TSU.MI.KI "magical cubes". TSU.MI.KI extracts in real-time the shapes the user attempts to construct, for example when confronted with a river the user can build a bridge, a ship or a plane from the cubes and the system will follow with a matching virtual plane model and a corresponding narrative in the story space. Furthermore, in TSU.MI.KI the physical set of ActiveCubes with its sensors and actuators becomes physical controllers of virtual story entities, for example when the user manipulates the physical plane she constructed the virtual plane will follow its tracked movements while the physical plane vibrates and rotates its propeller.

To date, the use of the ActiveCube system as a concrete, ecologically valid tool to understand dynamic functional processes underlying constructional ability in either typically developed children or in children with neurological pathology has not yet been explored. Developmental Coordination Disorder (DCD) is a marked impairment in the development of motor coordination that significantly interferes with academic achievement or activities of daily living (American Psychiatric Association, 1994). The motor coordination of children with DCD is substantially below that expected given the child's chronological age and measured intelligence. Older children may display difficulties with the motor aspects of assembling puzzles, building models, playing ball, and printing or writing. The task of construction embraces two broad classes of activities: drawing and building or assembling (Fisher & Loring, 2004). Constructional activity entails spatial perception with a motor response (Fisher & Loring, 2004).

Visuospatial constructional ability is complex, comprising multiple, distinct, but interrelated subcomponents (Cronin-Galomb & Braun, 1997). These include the ability to combine elements into meaningful wholes, to discriminate between objects, to distinguish between left and right, to understand relationships among objects in space, to adopt various perspectives in order to represent and rotate objects mentally, to comprehend and interpret symbolic representations of external space, and to work out solutions for non-verbal problems (Fernando, et al, 2003). It is essential to assess and treat clients with visuospatial deficits because constructional deficits have been shown to be related to poor activity in daily functioning (Katz et al, 1997).

The overall goal of the present research is to develop and evaluate a paradigm in which TUIs are used for the study, assessment and intervention of dynamic constructional processes among typically developed children and those with DCD. The objective of this paper is to describe the ActiveCube system and the special interface designed for assessing and treating children with DCD. The specific objectives of this present pilot study were: (1) to examine the feasibility and the usability of the Active Cube TUI in children with DCD as well as in typically developed children in terms of their ability to manipulate ActiveCube, their level of enjoyment and their extent efforts in the task, (2) to examine whether this novel interface distinguishes between children with DCD with constructional deficits and typically developed children, and (3) to establish the protocol for future research with the system.

2. METHODS

2.1 Participants

The experimental group consisted of: three male children with DCD, aged 6 to 7 years, who were recruited from a local Child Development Clinic. A preliminary screening of these children as having DCD was based on parental responses to the Children's Activity Questionnaire for early identification of children who suspected as having DCD (Rosenblum, in press). These children were also reported by the Clinic's occupational therapists as having difficulties in copying and constructing models. The children received a score of 5% or below on the Movement Assessment Battery for Children (M-ABC) (Henderson & Sugen, 1992). This score indicated the presence of significant motor coordination difficulties. Finally, a clinical diagnosis of DCD was made by a neurologist based on standard criteria for DCD, as outlined by the Diagnostic and Statistics Manual, DSM IV (American Psychiatric Association, 1994).

The control group consisted of three typically developed male children, aged 6 to 7 years, who were recruited via a convenience sample. They had no motor deficiencies, as evaluated by the M-ABC and had no difficulties in copying geometric figures, as evaluated by the Developmental Test of Visual-Motor Integration (VMI) (Beery, 1997). In addition, for visuospatial organization , as evaluated by the subtest Block Design from the Wechsler Intelligence Scale for Children (WISC-R95) (Cahan, 1998), they received scores in the range of the average norm. The parents of all six children gave their consent to participation in the study.

2.2 Instruments

2.2.1 Hardware. The Active Cube system, shown in Figure 1, is an automated system for constructional cognitive assessment as described above.

2.2.2 Software. A playground metaphor, shown in Figure 2, is used to present the assessment and intervention tasks. The playground includes six 3D playground structures: seesaw, dog, airplane, slide, carrousel and pyramid. Each apparatus may be constructed from up to seven cubes. The maximum number of cubes was reduced to seven from an initial ten cubes following a preliminary study that demonstrated that the combined mass of cubes makes larger structures unstable and cumbersome. Once an apparatus is selected, it is enlarged and presented on the screen by itself. The child's task was to successively use the ActiveCube system to construct this apparatus in a "matching" strategy that we have termed "Perspective Matching" (see Figure 3). In this strategy, the designated apparatus is displayed in its entirety; the child views only a perspective image of the structure or the screen. The child's task is to construct the displayed structure using the cubes.

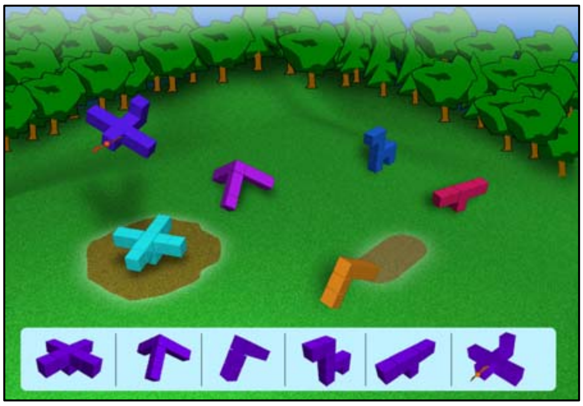


Figure 2. *The playground metaphor that is used to present the assessment and intervention tasks. Each of the six playground apparatus may be constructed from up to seven cubes.*

Figure 3. *One of the children who participated in the pilot study constructing an apparatus from the ActiveCube.*

A second construction strategy, "Following" was provided for children who were unable to correctly complete the structure with "Perspective Matching". In this case, the child constructs the designated apparatus, cube by cube, following a set of successive steps displayed on the screen. First, the base cube appears onscreen. Next, a second cube, which needs to be connected to the base cube, appears and flashes on and off on the screen. As soon as the child attaches this cube correctly, a third cube appears. The process continues until the structure is completed correctly.

2.2.3 Clinical assessment tools. The Movement Assessment Battery for Children (M-ABC) and DCD questionnaire were used to identify whether the participants met the diagnostic criteria for DCD. The Developmental Test of Visual-Motor Integration (VMI), Block Design test were performed to all participants to detect constructional deficits. The M-ABC was developed by Henderson & Sugen (1992) for children, aged 4-12 years with the aim of measuring the level of motor functioning; it has been demonstrated to have good test-retest and inter-rater reliability, and concurrent validity. The VMI was developed by Beery (1997) to test children aged 3-17 years with the aim of evaluating visual-motor integration; it has predictive validity for learning difficulties in various areas, particularly in writing and reading. The Children Activity Scale for Parents (ChAS-P) was developed by Rosenblum (in press) to identify children with Developmental Coordination Disorders aged 4-8 years. Internal consistency and content, construct and face validity have been established as well as a cut-off scores. The Wechsler Intelligence Scale for Children (WISC), is an intelligence test for children. We used the third edition of the Hebrew version of the WISC called WISC-R95, with Israeli norms for children ages 6-14 years. We used the WISC's subtest: Block Design, a construction test. In this WISC subtest children use red-and-white blocks to construct a pattern according to a displayed model.

2.2.4 Outcome measures. All sessions were recorded with a digital video recorder for subsequent viewing and analysis. A 3-point "Enjoyment Scale" was used to query each of the children following the Perspective Matching session. They were asked to rate their level of enjoyment while using the ActiveCube constructional activities. A 3-point "Effort Scale" was also used to query the amount of effort perceived by each child while performing each task.

All connections of the ActiveCube system were recorded online for subsequent offline analysis. The following outcomes were calculated:

Onset time − time from the start of the task (when the designated structure is displayed on the screen) until the first cube is connected to the base cube.

Connect time − time taken for each cube to be connected to or disconnected from the base unit.

Total time − time from the beginning of the task (when the designated structure is displayed on the screen) to its completion (when the last cube has been connected or disconnected).

Similarity − the similarity between the designated apparatus as displayed on the screen and the apparatus as constructed by the child is used as a measure of accuracy. When the structure is not identical to the original apparatus, a score is automatically computed by the system according to the number of cubes that have been connected correctly and the number that have not been connected correctly. Equation 1 for measuring similarity is the one described by Sharlin et al. (2002) where *i* is an intersection of *s* (the structure which the child built) and *p* (the prototype), and $|i|$, $|s|$, and $|p|$ are the number of cubes in *i*, *s* and *p* is maximized over all possible intersections *i* produced by rotating or translating *s*. Intuitively speaking, the algorithm computes similarity by the number of intersecting cubes minus the number of remaining "extra" cubes in the participant's structure, normalized by the number of cubes in the prototype.

$$Sim = 100 \frac{|i|}{\max(|p|,|s|)} \tag{1}$$

The similarity function serves two main purposes: (1) to provide real-time feedback regarding the integrity of each ActiveCube connection and (2) to automatically compute the number of errors where error is defined as a connection or a disconnection which decreases the similarity and thus does not progress the task.

2.3 Procedure

The research conducted in a quiet environment (a special kindergarten room was used for the control group and a secluded room at the clinic was used for the experimental group). The children sat at a table suited to their anthropometric characteristics. The ActiveCube system and construction tasks were described to each participant at the start of the session and specific instructions in accordance with the experimental protocol were provided. At the beginning of the session there was a practice task. During this task the child learned how to connect the cubes via demonstration. The child was then asked to connect a few cubes independently, receiving help if necessary. The child then constructed each of the six playground apparatus using the "Perspective Matching" strategy. The order of the six apparatus was predetermined based on their complexity with the easiest apparatus first. The children were instructed to press the Escape key when they had completed construction of each apparatus. The session lasted for about 30 minutes. The children who were unable to complete an apparatus using the "Perspective Matching" strategy (i.e., they could not achieve a similarity score of 100%), were requested to complete it using the "Following" strategy.

2.4 Data analysis

The video recording observations of the ActiveCube construction process was used to verify that all ActiveCube connections and disconnections were correctly recorded as such by the software. Due to small sample size, data analysis consisted primarily of descriptive statistics as well as the graphical comparison of data from individual participants.

3. RESULTS

3.1 Feasibility and usability of the ActiveCubes

All the children rated their use of the ActiveCubes to construct the various structures with the highest score ('enjoyed very much') of Enjoyment Scale. In addition, based on observation of the children while performing these tasks, it was evident that they were motivated and maintained a high level of alertness while using the ActiveCubes. None of the children indicated that they wanted to stop prior to completion of the structures and they all performed the tasks in a systematic and steady manner. The majority of the participants recognized the natural playground setting in which the six structures were located. They also identified each of the structures.

Fifty percent of the children reported that the task was easy; 33%of the children reported that the task was moderate (neither too difficult nor too easy) and 17%of the children reported that the task was difficult.

All of the children succeeded in connecting and disconnecting the ActiveCubes by the male/female metal press-stub connectors. In some cases, for technical reasons, although a connection appeared to be intact mechanically, it was not recorded by the computer as being connected electronically. Unstable mechanical connections also occasionally led to a collapse of the ActiveCube structure when the child exerted too much force.

3.2 Similarity scores for a child with DCD compared to a typically developed child

These pilot results show that the ActiveCube paradigm appears to be able to distinguish between children with DCD who had constructional difficulties and those who are typically developed. For example, the similarity scores achieved by two children using the Perspective Matching strategy are compared in Figure 4. The typically developed child (shown in light grey) achieved a score of 100% for all six apparatus (i.e., he succeeded in correctly constructing each of the playground structures). In contrast, the child with DCD (shown in dark grey) achieved scores ranging from 42% to 83% for the same tasks. Note that the level of difficulty of the structures varied. The child with DCD achieved similarity scores of about 40% for the airplane and pyramid apparatus, about 60% for the dog apparatus and over 80% for the seesaw, slide and carousel apparatus.

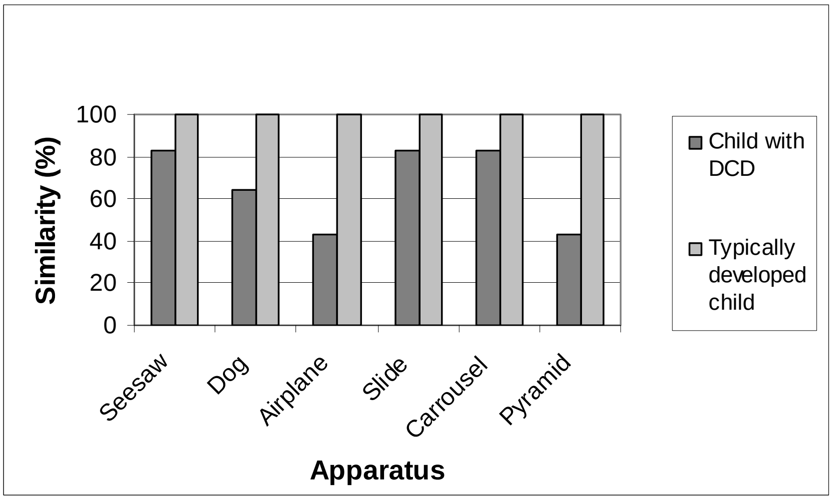


Figure 4. *Similarity scores for each of six apparatus for a typically developed child (light grey) and of a child with DCD (dark grey).*

In addition to providing a summary score for similarity, ActiveCube provides the ability to quantitatively track the similarity during the construction process. Figures 5 and 6 illustrate the process whereby various apparatus are constructed when using the Perspective Matching strategy. The similarity scores for a typically developed child who constructed the airplane (to be constructed from 7 cubes) are plotted as a function of time in Figure 5 (left panel). Figure 5 (right panel) shows a similar graph for the same structure for a child with DCD. Several differences in performance for the two children are apparent from a comparison of these two graphs. The typically developed child achieved a similarity score of 100% (perfect accuracy) in a 6-step task that took 130 s to complete. In contrast, the maximum similarity for the child with DCD was only 42%; he took 53 s to perform the task and completed only three steps.

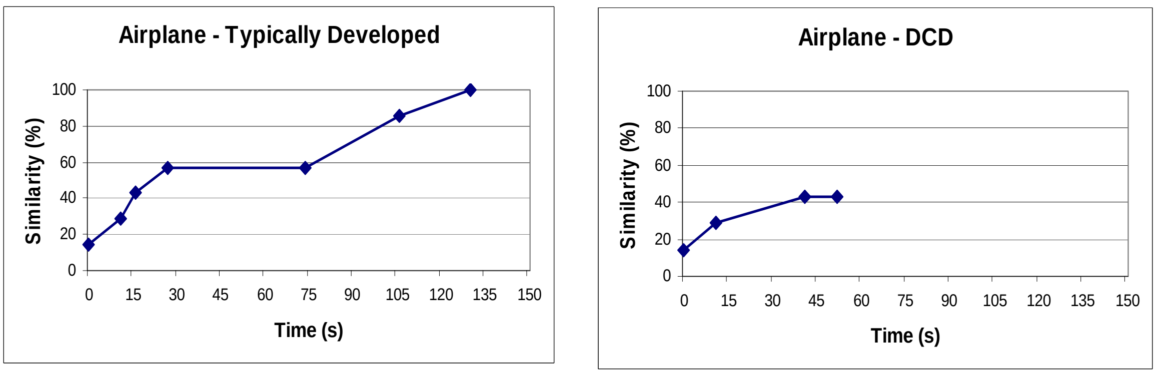
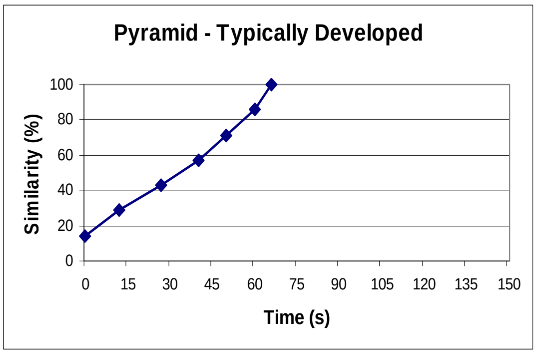


Figure 5. *Similarity versus time graph during the construction of the airplane apparatus by a typically developed child (left) and a child with DCD (right).*

The results for the pyramid apparatus (to be constructed from 7 cubes) are similar to those for the airplane apparatus. The similarity scores for a typically developed child who constructed the pyramid (to be constructed from 7 cubes) are plotted as a function of time in Figure 6 (left panel). Figure 6 (right panel) shows a similar graph for the same structure for a child with DCD. Again we note that the typically developed child achieved a similarity score of 100% (perfect accuracy) in a 6-step task that took 66s to complete. In contrast, the maximum similarity for the child with DCD was only 42%; he took 43 s to perform the task and completed only three steps.



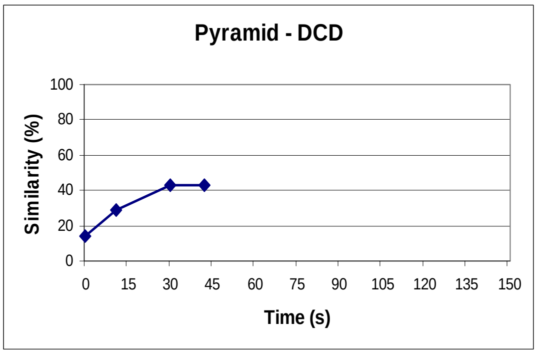


Figure 6. *Similarity versus time graph during the construction of the pyramid apparatus by a typically developed child (left) and a child with DCD (right).*

The child with DCD whose data are shown in Figures 5 and 6 was unable to complete the playground structures when using the Perspective Matching strategy. He was then given an opportunity to use the easier Following strategy wherein the child constructs the designated apparatus, cube by cube, following the successive steps as they are displayed on the screen. As shown in Figure 7, this child was able to correctly construct the airplane apparatus when using this strategy. He completed it in 151 s using six steps.

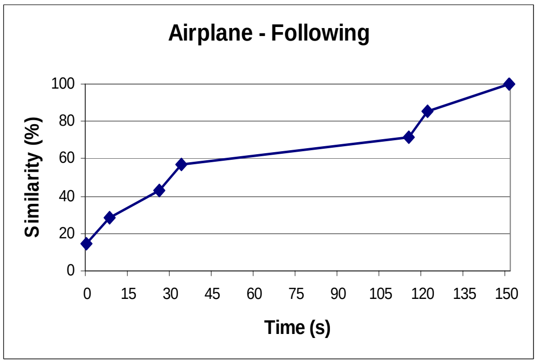


Figure 7. *Similarity versus time graph during the construction of the airplane apparatus by a child with DCD using the Following strategy.*

4. SUMMARY AND CONCLUSIONS

The results of this pilot study were highly positive in terms of verifying the participants' interest and enjoyment of the ActiveCube playground structure paradigm. They understood the tasks and were motivated to carry them out. A number of technical problems arose, most notably related to some difficulties in maintaining an intact connection between adjacent cubes. Another difficulty related to the total physical mass of the completed structure.

These results have confirmed several of the experimental protocol decisions. For example, based on our preliminary experience, we had chosen to limit the total number of cubes to no more than seven per apparatus during this pilot study. This number allowed the construction of structures of sufficient complexity and yet prevented the creation of masses that would lead to their collapse or become too wieldy for the children to manipulate during the post-construction play period. Despite the 7-cube limit, the structures appear to be sufficiently diverse in their levels of difficulty as demonstrated by the results shown in Figure 4.

On the other hand, the results have also provided important information concerning changes that will be needed prior to conducting the full study. The most important of these changes concerns the need to alter the mechanical method by which the cubes were attached to each other. Currently, adjacent cubes are attached by means of pressing their faces together; each cube contains two male and two female metal press-studs (one in each corner) which snap together. As indicated above, some connections appeared to the child to be intact, and yet were not recorded as such by the system. Moreover, the structures would sometimes collapse due to inadvertent disconnections between one or more cubes. The Human Interface Engineering Laboratory at the University of Osaka has now developed a newer version of the ActiveCube system which uses magnets (rather than press-stubs) for the inter-cube connections. Brief initial testing with a trial set of the magnet-based cubes has provided encouraging results regarding the integrity and stability of this alternate mechanism.

These initial results have provided important information concerning the feasibility and usability of the ActiveCube system for assessment and treatment of children with DCD. Despite the technical problems, the initial results show that the ActiveCube system and playground apparatus tasks appear to be sensitive to differences in the constructional abilities of the children.

Acknowledgements: We thank Galia Gutwil for her help with participant recruitment and data collection. This research was supported in part by Special Coordination Funds of the Science and Technology Agency of the Japanese Government, by a Canadian NSERC Discovery Grant, and by the University of Haifa Development Fund.

5. REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders,* 4th ed., American Psychiatric Association, Washington, D.C.

K E Beery (1997). *The Beery-Buktenica Developmental Test of Visual-Motor Integration,* 4th ed., Modern Curriculum Press, New Jersey.

S Cahan (1998), Manual for the WISC-R95. Ministry of Education, The Psychological and Counseling Service, Jerusalem, Israel.

A Cronin-Golomb and A E Braun (1997), Visuospatial dysfunction and problem solving in Parkinson's disease. *Neuropsychology*, **11**, *1*, pp. 44-52.

J S Fisher and D W Loring (2004). Construction. In *Neuropsychological Assessment,* 4th ed., (M. D. Lezak, D. B. Howieson, & D. W. Loring , Eds), Oxford University Press, New York, pp. 531-568.

M Gorbet, M Orth and H Ishii (1998), Triangles: Tangible interface for manipulation and exploration of digital information topography, *Conference on Human Factors in Computing Systems,* pp. 49-56.

S E Henderson and D A Sugen (1992), *Movement assessment battery for children*. Psychological Corporation/Harcourt Brace-Jovanovich, New York.

Y Itoh, S Akinobu, H Ichida, R Watanabe, Y Kitamura and F Kishino (2004), TSU.MI.KI: Stimulating Children's Creativity and Imagination with Interactive Blocks. *Second International Conference on Creating, Connecting and Collaborating through Computing*, pp. 62-70.

N Katz, D Champagne and S Cermak (1997), Comparison of the performance of younger and older adults on three versions of a puzzle reproduction task. *American Journal of Occupational Therapy*, **51**, *7*, pp.562-8.

Y Kitamura, Y Itoh and F Kishino (2001), Real Time Interaction with ActiveCube, *Conference on Human Factors in Computing Systems,* pp. 355-356.

S Rosenblum (in press). The development and standardization of the Children Activity Scale (ChAS-P/T) for the early identification of children with Developmental Coordination Disorders (DCD). *Child Care Health and Development*.

E Sharlin, Y Itoh, B Watson, Y Kitamura, S Sutphen and L Liu (2002), Cognitive Cubes: A Tangible User Interface for Cognitive Assessment, *Conference on Human Factors in Computing Systems*, pp. 347-354.

E Sharlin, Y Itoh, B Watson, Y Kitamura, S Sutphen, L Liu and F. Kishino (2004), On tangible user interfaces, humans and spatiality, *Journal of Personal and Ubiquitous Computing*, **8**, *5*, pp. 338-346.

B Ullmer and H Ishii (2001), Emerging Frameworks for Tangible User Interfaces. In Human Computer Interaction in the New Millennium, J. M. Carroll, Ed. Addison-Wesley, pp. 579-601.

R Watanabe, Y Itoh, M Kawai, Y Kitamura, F Kishino and H Kikuchi. Implementation of ActiveCube as an Intuitive 3D Computer Interface, Proc. of 4th International Symposium on Smart Graphics, pp. 43-53.

An evaluation of the use of a switch controlled computer game in improving the choice reaction time of adults with intellectual disabilities

P J Standen[1], R Karsandas[2], N Anderton[3], S Battersby[4] and D J Brown[5]

[1,2,3]Division of Rehabilitation & Ageing, University of Nottingham,
B Floor, QMC, Clifton Boulevard, Nottingham NG7 2UH

[4,5]School of Computing & Mathematics, Nottingham Trent University,
Newton Building, Burton Street, Nottingham NG1 4UB

p.standen@nottingham.ac.uk, rajkarsandas@hotmail.com, nicola.anderton@nottingham.ac.uk, stephen.battersby@ntu.ac.uk, david.brown@ntu.ac.uk

[1,2,3]*www.nottingham.ac.uk/rehab*, [4,5]*www.isrg.org.uk*

ABSTRACT

The inability of people with intellectual disabilities to make choices may result from their lack of opportunities to practice this skill. Interactive software may provide these opportunities and software that requires a timed response may reduce choice reaction time. To test this, 16 people with severe intellectual disabilities were randomly allocated to either an intervention or a control group. The intervention group spent eight sessions playing a switch controlled computer game that required a timed response while the control group spent the same amount of time playing a computer based matching game that did not require a timed response. Both groups repeated a test of choice reaction time (CRT) that they had completed prior to the intervention. The intervention group made more accurate switch presses with repeated sessions while receiving less help from the tutor who sat alongside them. The intervention group also showed a significant reduction in their CRT from baseline while the control group did not.

1. INTRODUCTION

People with intellectual disabilities are amongst the most socially excluded and vulnerable groups in many countries and the intention of current policy (eg Department of Health, 2001) is to enable them to have as much choice and control as possible over their lives, be involved in their communities and to make a valued contribution to the world at work.

However, research suggests that they have difficulty making choices or decisions. For example, Jenkinson and Nelms (1994) found that they were less likely to produce responses that reflected a systematic evaluation of possible alternatives than were people without intellectual disabilities. Although the terms “decision” and “choice” making are quite often used interchangeably in the literature, Mincemoyer and Perkins (2003) describe choice making as selecting from among identified alternatives and only one component of decision making. Similarly, Jenkinson and Nelms (1994) state that being able to choose among options is a necessary prerequisite to decision-making and that decision making involves more than a simple expression of preferences. It involves an understanding of an issue, identification and informed evaluation of options, communications of a decision and commitment to action. Distinguishing the terms in this way is helpful and underlines the benefits that might follow from promoting choice making.

Two areas of research suggest that there may be potential for promoting the ability to make choices in people with intellectual disabilities. First, Cooper & Browder (2001) see their inability to make choices as resulting from a constant denial of choice. This leads to an inability to respond to stimuli in the community and as a result they may never acquire skills that enhance their independence. Research examining opportunities for choice making consistently identifies ability as a determinant (eg Stancliffe et al, 2000). However, reduced intellectual capacity need not necessarily lead to the inability to make choices as even people with severe and complex disabilities can express stable preferences when provided with choices (Lancioni et al, 1996). Cooper & Browder (2001) demonstrated how embedding choice opportunities in a trip to a fast food outlet can enhance choice making. They set out to increase the opportunities for choice making

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

in a group of eight people with severe intellectual disabilities with a staff training package. They trained staff to offer a series of two options (eg which of two doors through which to enter; two photos of food or drink options) when using fast food restaurants. Compared with baseline, the people with intellectual disabilities increased the number of choice responses they made both prompted and independent. They also required a much lower level of prompting to make these choices after the intervention.

The implications of this study are that carers of those with intellectual disabilities need to present them with options on as many occasions as possible. Options presented by carers in day to day activities may well be the most ecologically valid way of increasing the opportunities for making decisions. However, they are often pressed for time, wary of letting their charges take risks and may also find it difficult to suppress their natural inclination to take over before allowing the person sufficient prompts for them to perform the selection independently. The second area of research suggests that interactive computer software should be explored as a low risk means of supplementing these limited opportunities.

There is now an increasing body of work exploring the effects of interactive computer software on cognition. Green & Bavelier (2003) found that computer games requiring frequent switches of attention had a beneficial effect on visual attention that generalised to a different task in non-disabled young people. In a review of both the positive and negative effects of playing videogames, Griffiths (2004) describes the role of videogames in cognitive rehabilitation for example in perceptual disorders, conceptual thinking, attention, concentration and memory in patients with brain damage following stroke or trauma. Claims have also been made for the beneficial effects on cognitive skills of virtual environments (VE). Preliminary work suggests they have a role in improving memory (Rose et al, 1999) and attention (Rizzo et al, 2000) in those recovering from brain injury and other neurological disorders.

One easily measured aspect of choice making is reaction time. To investigate whether this improved after use of VE by people with severe intellectual disabilities, Standen & Ip (2002) devised two tests to measure choice reaction time at baseline and post intervention. The first consisted of 10 cards depicting a familiar object on one side that were shuffled and placed face down before the participant was asked to pick one depicting a particular object eg apple. Time taken to do each of 10 trials was recorded. The second task involved the participant choosing two items from each of 10 pictorial shopping lists displayed on a computer monitor. Again, time taken was recorded. To avoid habitual responses, on three trials, after the time for the second choice had been recorded, participants were told that one of the items was no longer available and they had to choose something else. Data from 9 individuals who had six sessions using virtual environments (the active group) indicated a significant ($p<0.003$) reduction from baseline in their time to make a choice in the card game and the shopping list. Although there was some improvement when they were forced to make an alternative choice in the shopping list this reduction did not reach significance. A control group of 6 individuals who sat alongside a matched active partner during the intervention showed no improvement on all three measures.

Although these results are encouraging, the outcome measures for this study can be criticised. The shopping list task was computer based and any improvement on this task may have merely demonstrated increasing familiarity with using a computer. For the non-computer based outcome measure (the card game) timing was dependent on the reaction time of the research assistant who used a stopwatch to record time to make a choice. In addition, members of the control group were reluctant to continue with the study even though they were promised sessions using the VE after the study had finished. The current study was designed to avoid these drawbacks by using a non-computer based outcome measure with accurate electronic timing and by giving the control group a task to complete. In order to involve participants who had difficulty controlling the interfaces for the three dimensional VE, the intervention was a game that could be controlled by a switch. As a more disabled group was taking part the first aim of the study was to see whether they could learn to play the game The second aim was to see whether repeated sessions playing the game resulted in a decrease in choice reaction time (CRT).

2. METHODS

2.1 Design

To test the first aim, for the intervention group the change over repeated sessions in the proportion of switch presses made that were made at the right time (correct switch presses) was compared with chance level. The proportion of correct switch presses was correlated with help received to determine whether switch pressing was made independently. To test the second aim, baseline CRT was compared with post intervention CRT in the intervention group and a matched control group.

2.2 Participants

24 people from a day centre for adults with intellectual disabilities were nominated by their specialist carers at the centre if they matched the inclusion criteria below:

1. Severe intellectual disabilities indicated by their scores on the Adaptive Behaviour Scale (ABS) (Nihira, Leyland and Lambert, 1993).
2. No visual impairment, which prevented viewing the screen
3. Sufficient motor and cognitive ability to operate a simple microswitch
4. No previous exposure to microswitches or similar software

Only 16 of those nominated could be included in the study as 8 were unable to react within the recording limit of 1750 milliseconds on a minimum of 5 of the 20 baseline choice reaction trials. The remaining participants were grouped into pairs matched in terms of their age, ABS score and mean baseline CRT and members of each pair were randomly allocated to either the intervention or the control group. Their characteristics are shown in Table 1. Although the mean CRT is longer for the control group this difference is not statistically significant.

Table 1. *Characteristics of participants.*

	Intervention Group n=8	**Control Group n=8**
Mean (SD) Age in Years	37.5 (15)	35.5 (8)
Mean (SD) ABS scores	37.1 (8.6)	37.0 (7.7)
Mean (SD) baseline CRT in milliseconds	1272.7 (247.3)	1352.9 (231.7)
Male: Female Ratio	5:3	4:4
Ethnicity	7 Caucasian:1 Asian	7 Caucasian:1 Asian

2.3 Interventions

The computer game was designed specifically for people with severe to profound intellectual disabilities by the fourth author. It involved a man jogging across the screen and encountering several obstacles such as a log or a large rock. A screen shot from the game is shown in Figure 1. The aim was to jump over these hurdles and to make it to the finish line when audio (a cheering sound) as well as visual (a rotating trophy) signals appeared as a reward. The man was made to jump by a single press of a 'jelly bean' microswitch which if made at the appropriate time would allow the man to clear the obstacle. However if the switch was continuously pressed, there was no observed effect within the game. The game recorded the number of switch presses that resulted in the running man clearing the obstacle and the total number of switch presses made.

The control task involved the participant trying to match pictures of familiar animals or objects from a series of cards to those displayed on a computer monitor. The participant sat in front of the monitor and ten cards faced down were placed on the table in front of her. On the other side of each card was a different picture depicting an object familiar to the participant (eg animals, food). When a picture appeared on the monitor the participant had to find the card with the matching picture. The rate at which the pictures appeared on the computer monitor was controlled by the researcher and a new picture only appeared once the participant had found the card with the matching picture.

2.4 Outcome measures

Choice reaction time. This device, designed by Piercy Adams Ltd. Cambridge, involved a series of 6 outer red LEDs as well as a central green LED, all of which were placed within a close proximity to their respective light sensitive diodes (see Figure 2). The device measured choice reaction time by assessing how quickly a user could shift his/her hand from the central green LED to the appropriate outer red LED which became illuminated at random. A new light was only illuminated once the participant had responded to the previous one. However the timer would only record choice reaction times with a maximum duration of 1750 milliseconds and therefore any reaction time above this was not recorded but noted as the maximum.

Help received. All sessions with the intervention group were videotaped and analysed using OBSWIN (Oliver et al, 2000) to measure amount of tutor assistance (measured as an event variable). The definition of assistance given was taken from a methodology previously established by Standen et al (2002). It referred to any help given to the participant to carry out a correct switch press or to direct the participant's hand closer to the switch.

Figure 1. *Screen shot from computer game.*

Figure 2. *The Piercy Adams Choice Reaction Timer.*

2.5 Data collection

Data collection took place in the day centre attended by the participants. And all data collection was carried out by the same member of the research team (RK).

Measuring CRT Before baseline measures were recorded the test was demonstrated to the participant and a verbal explanation given if they could understand it. They were then encouraged to complete 20 trials and if they could manage at least 5 they went on to be included in the study. At post intervention, they received another demonstration from the researcher if it was required before being asked to complete another 20 trials.

The intervention Each participant spent 8, twice weekly sessions that lasted up to a maximum of 10 minutes but could be terminated earlier if the participant did not wish to continue.. This session length was kept the same for intervention and control participants of each matched pair. One of the research team (RK) sat alongside the participant to give them assistance and encouragement.

2.6 Analysis

As the upper limit for a reaction time was set at 1750 milliseconds, for each participant the median of their CRTs for each 20 trial block was calculated. As these values were normally distributed the comparison between their baseline and post intervention median CRT was made using a paired t-test. To take account of any variation in timer playing the game and a strategy of high levels of switch pressing, the number of correct switch presses from the computer game was converted to a proportion of the total number of switch presses made. All videotapes were analysed by the same researcher who had collected the data so in an attempt to minimise bias resulting from knowledge of how many sessions the participant had completed, the order of the tapes for analysis was scrambled. Video collected data were expressed as a percentage of session duration. Comparisons between first and last sessions for both correct switch presses and help received were made using a paired t-test. The relationship between these two variables was examined using Pearson's correlation coefficient. The comparison of the change in the proportion of correct switch presses with chance level was made using regression analysis.

3. RESULTS

3.1 Change in the proportion of accurate switch presses with increasing sessions

The proportions of accurate switch presses by the intervention group increased significantly ($p<0.0001$) from first to last session. Regression analysis produced a significant ($p<0.00003$) adjusted R squared value of 0.956 indicating that over 95% of the increase in proportion of accurate switch presses was explained by session number (i.e. increased exposure to the game). A beta value of 0.038 ($p<0.00002$), indicates that with each session, the proportion of accurate switch presses increases by approximately 0.038. Figure 3 shows the group means for each of the eight sessions.

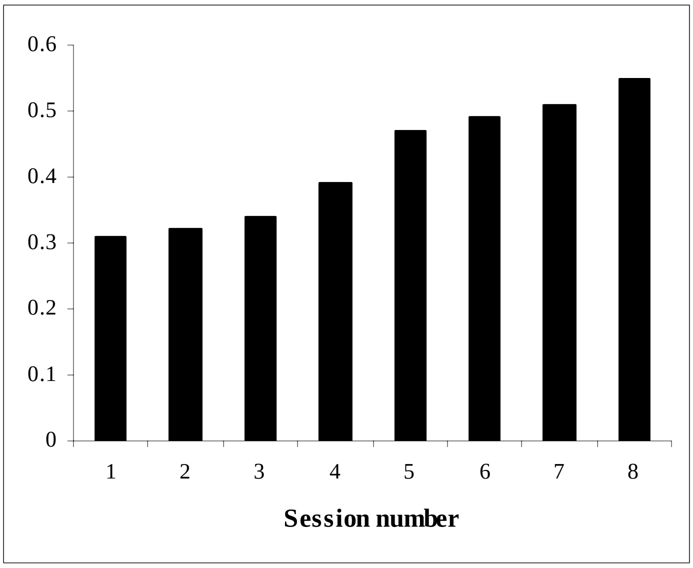


Figure 3. *Mean proportions of accurate switch presses by session.*

The increase in the proportion of accurate switch presses was not dependent on help from the tutor as the percentage of the session in which participants received help from the tutor decreased significantly ($p<0.0002$) from sessions 1 to 8 and was significantly ($p<0.00007$) negatively correlated with the proportion of accurate switch presses made increased. Figure 4 shows the group means for each of the eight sessions.

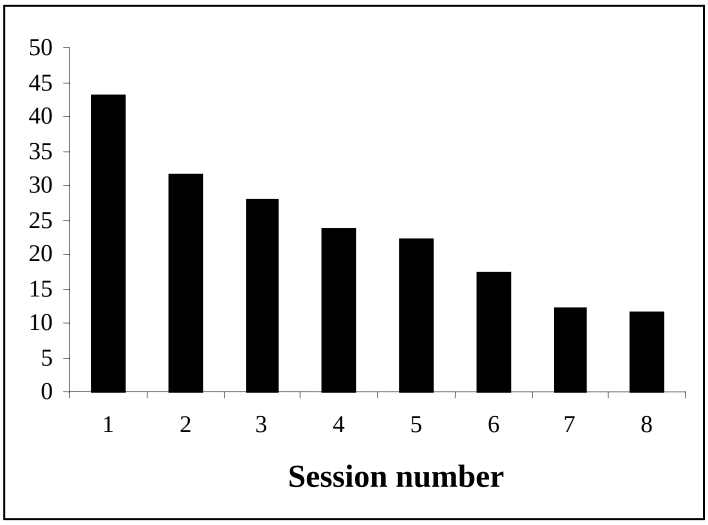


Figure 4. *Mean percentage of sessions spent receiving physical assistance*

3.2 Change in CRT from baseline to post intervention

Mean choice reaction times are shown in Table 2. The intervention group showed a significant ($p<0.003$) decrease from baseline to post intervention in mean choice reaction time whereas the decrease observed in the control group failed to reach significance.

Table 2. *Baseline and Post intervention mean Choice Reaction Times in milliseconds.*

	Intervention Group		Control group	
	Baseline	Post intervention	Baseline	Post intervention
Mean CRT (msec) (SD)	1273 (247)	990 (170)	1353 (232)	1302 (229)

4. DISCUSSION

Even with the limited time playing the switch controlled game, the intervention group learnt how to play as evidenced by the increase in successful switch presses expressed as a proportion of the total number of switch presses. From this we can conclude that the design of the game matched their level of ability and that their interest and attention could be sustained over eight sessions. As with all of our research, someone sat alongside the participant and there is always the possibility that the researcher or tutor can cross the fine line between facilitating the participant’s performance and controlling it. In initial sessions the participants will inevitably require high levels of help from the tutor but as they become more familiar with the game and more skilled the level of help from the tutor should diminish. The results indicate that this was happening so it can be concluded that the increasing proportions of correct switch pressing are evidence of learning or increased ability in the participant. For a group of people who may experience very low levels of stimulation or social interaction such an activity can have a range of benefits. According to Lancioni et al. (2001) who reviewed several studies on the use of microswitches “the establishment of successful responding may have immediate benefit in terms of increased alertness and positive interaction with the surrounding world”.

The intention of this study was to explore whether any other benefits could be detected from the continued use of a switch in order to play a computer game. When compared to the control group who spent the same amount of time playing a game without a time limited response, the intervention group showed a significant decrease in choice reaction time after eight sessions playing the game and this in spite of small sample sizes. Both groups had equal lengths of time sitting in front of the computer screen and similar rates of tutor interaction, suggesting that the time pressured nature of the task faced by the intervention group was the influential element.

There are obvious limitations in this study in addition to the small sample size. The sample was also fairly carefully selected so that there was limited variation in their ability so these findings may not generalise to other groups. The participants only received a limited exposure time to the intervention and the equipment to measure CRT limited the type of participants we could involve. Although measures were taken to avoid researcher bias there is always the possibility that knowledge of which group the participant belonged to might have subtly influenced the encouragement given by the researcher when CRT was measured post intervention. Unlike non-disabled groups, these participants cannot be left alone in a separate room to minimise extraneous influence when using such equipment.

Even if these finding are robust and would be repeated with a larger sample size, questions still remain. Can we identify groups who would benefit from such an experience while others would not? Does improving CRT actually facilitate choice making as in the selection from options? Does this ability facilitate decision making as suggested by Jenkinson and Nelms (1994)? While using interactive software in the form of computer games may well improve alertness and responsiveness, the facilitation of the selection from among options or decision making may require more complex games or virtual environments. By depicting more complex every day situations such software might be used to facilitate other components of the decision making process such as the understanding of an issue and the identification and informed evaluation of options which Jenkinson and Nelms (1994) consider essential parts of the process. This software could also be used to depict situations in which participants can practice these skills.

5. REFERENCES

K J Cooper, D M Browder DM (2001), Preparing staff to enhance active participation of adults with severe disabilities by offering choice and prompting performance during a community purchasing activity, *Research in Developmental Disabilities,* **22**, pp.1-20.

Department of Health (2001), Valuing People: a New Strategy for Learning Disability for the 21st Century. HMSO: London.

C S Green, D Bavelier. (2003), Action video game modifies visual selective attention,. *Nature,* **423**, pp. 534-37.

M Griffiths (2004), Can videogames be good for your health? *Journal of Health Psychology,* **9**, 3, pp.

339-344

J Jenkinson, R Nelms (1994), Patterns of decision making behaviour by people with intellectual disability: an exploratory study, *Australia and New Zealand Journal of Developmental Disabilities*, **19**, 2, pp. 99-109.

G E Lancioni , M F O'Reilly, E Emerson (1996), A review of choice research with people with severe and profound disabilities. *Research in Developmental Disabilities,* **17**, pp. 391-411.

G E Lancioni , M F O'Reilly, G Basili (2001) Use of Microswitches and speech output systems with people with severe/profound or multiple disabilities: A literature review, *Research in Developmental Disabilities*, **22**, pp. 21-40

C C Mincemoyer, D F Perkins (2003), Assessing decision making skills of youth, *Forum for Family and Consumer Issues,* **8**, 1, http://www.ces.ncsu.edu/depts/fcs/pub/8(1)/mincemoyer.html

K Nihira K, Leland H, Lambert N. *Adaptive Behaviour Scale – Residential and Community* (second edition). Washington, DC: American Association on Mental Retardation,1998.

C Oliver, N Martin, S Hall (2000) OBSWIN-Observational Data Collection & Analysis, Antman Ltd, London

A A Rizzo, J G Buckwalter, T Bowerley, A Van Rooyen, J S McGee, C Van der Zaag, U Neumann, M Thiebaux, L Kim, C Chua (2000), Virtual environment applications for the assessment and rehabilitation of attention and visuospatial cognitive processes: an update. In: Sharkey PM, Cesarini A, Pugnetti L, Rizzo A, eds. *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies.* pp.197-207.

F D Rose, B M Brooks, E A Attree, D M Parslow, A G Leadbetter, J E McNeil, S Jayawardena, R Greenwood, J Potter (1999), A preliminary investigation into the use of virtual environments in memory retraining after vascular brain injury: indications for future strategy? *Disability and Rehabilitation,* **21**, pp. 548-554.

R J Stancliffe, B Abery, J Smith (2000), Personal control and the ecology of community living settings: beyond living-unit size and type,. *American Journal on Mental Retardation,* **105**, pp. 431- 454.

P J Standen, W M D Ip (2002), An evaluation of the use of virtual environments in improving choice reaction time in people with severe intellectual disabilities. In: Sharkey PM, Sik Lányi C, Standen PJ, eds. *Proceedings of the Fourth International Conference on Disability, Virtual Reality and Associated Technologies.* pp. 19-24.

P J Standen, D J Brown, M Horan, T Proctor (2002) How tutors assist adults with learning disabilities to use virtual environments, *Disability and Rehabilitation,* **24**, 11-12, pp. 570-577.

ICDVRAT 2006

Session IX. Stroke Rehabilitation

Chair: Sue Cobb

Brain-computer music interface for generative music

E R Miranda

[1]Interdisciplinary Centre for Computer Music Research, University of Plymouth,
Drake Circus, Plymouth, UK

eduardo.miranda@plymouth.ac.uk

cmr.soc.plymouth.ac.uk

ABSTRACT

This paper introduces a brain-computer interface (BCI) system that uses electroencephalogram (EEG) information to steer generative rules in order to compose and perform music. It starts by noting the various attempts at the design of BCI systems, including systems for music. Then it presents a short technical introduction to EEG sensing and analysis. Next, it introduces the generative music component of the system, which employs an Artificial Intelligence technique for the computer-replication of musical styles. The system constantly monitors the EEG of the subject and activates generative rules associated with the activity of different frequency bands of the spectrum of the EEG signal. The system also measures the complexity of the EEG signal in order to modulate the tempo (beat) and dynamics (loudness) of the performance.

1. INTRODUCTION

Research into Brain-Computer Interface (BCI) for music is an interesting arena for the development of new possibilities in recreational and therapeutic devices for people with physical and neurological disabilities; we shall refer to these systems as Brain-Computer Music Interfaces (BCMI). There are various music-making devices available for those with disabilities, and even though these devices have proved to work very effectively, they often do not allow as much control for those with severe physical disabilities. At present, access music tutors use gesture devices and adapted accessible technology to make this possible, which achieve excellent results; although for people with severe physical disabilities, having complete control of the environment created for them by the facilitator can sometimes prove difficult. This paper introduces a BCMI that uses the EEG steer generative musical rules in order to compose and perform music on a MIDI-controlled mechanical acoustic piano (Figure 1a).

Human brainwaves were first measured in 1924 by Hans Berger, who termed these measured brain electrical signals the electroencephalogram (literally "brain electricity writing"). Berger first published his brainwave results in 1929 (Berger 1929). Today, the EEG has become one of the most useful tools in the diagnosis of epilepsy and other neurological disorders. Further, the fact that a machine can read signals from the brain has sparked the imaginations of scientists, artists and other enthusiasts, and EEG has made its way into applications other than clinical. In the early 1970s, Jacques Vidal, a researcher at the University of California Los Angeles, did the first tentative work towards a brain-computer interface (BCI) system. The results of this work were published in 1973 (Vidal 1973). Despite a number of isolated initiatives at building BCI systems, research in this area did not take off until the early 1990s, probably due to technological limitations. In 1990, Jonathan Wolpaw and others at the Wadsworth Center at State University of New York Albany and Rensselaer Polytechnic Institute developed a system to allow some control of a computer cursor by individuals with severe motor deficits. Users were trained to use aspects of their EEG to move a cursor on a computer screen (Wolpaw et al. 1991). In 1998, Christoph Guger and Gert Pfurtscheller (Guger and Pfurtscheller 1998) presented a paper at the *9th European Congress of Clinical Neurophysiology* reporting impressive advances in BCI research: an EEG-based system to control a prosthetic hand. Many attempts followed with various degrees of success. For recent reports on BCI research please refer to the special issue of *IEEE Transactions on Biomedical Engineering* published in June 2004 (Vol. 51). To date, most efforts of BCI research have been aimed at developing assistive technology to help people communicate via computer systems and/or control mechanical tools, such as a wheelchair or a prosthetic organ. Comparatively little has been done to address the use of BCI technology for musical applications; such applications could undoubtedly enhance the quality of life for people with disabilities in many ways, be they recreational, therapeutic or professional.

As early as 1934, a paper in the journal *Brain* had reported a method to listen to the EEG (Adrian and Mathews 1934), but it is generally accepted that it was composer Alvin Lucier, in 1965, who composed the first musical piece using EEG: *Music for Solo Performer*, a piece for percussion instruments played by the resonance of the performer's EEG. Lucier placed electrodes on his own scalp, amplified the signals, and relayed them through loudspeakers that were directly coupled to percussion instruments, including large gongs, cymbals, tympani, metal ashcans, cardboard boxes, bass and snare drums (Lucier 1976). The low frequency vibrations emitted by the loudspeakers set the surfaces and membranes of the percussion instruments into vibration. Later in the 1960s, Richard Teitelbaum used various biological signals including the EEG and ECG (electrocardiogram) to control electronic synthesisers. In the early 1970s David Rosenboom (1975) began systematic research into the potential of EEG to generate art works, including music. Drawing on concepts from Cybernetics, he developed EEG-based musical interfaces associated with a number of compositional and performance environments that used the state-of-the-art EEG technology of the day. In particular, Rosenboom explored the hypothesis that it might be possible to detect certain aspects of our musical experience in the EEG signal. For example, in a paper that appeared in 1990 in *Computer Music Journal*, he introduced a musical system whose parameters were driven by EEG components believed to be associated with shifts of the performer's selective attention (Rosenboom 1990). Thirteen years later, this author and colleagues published a new paper in *Computer Music Journal* reporting experiments and techniques to enhance the EEG signal and train the computer to identify EEG patterns associated with different cognitive musical tasks (Miranda et al. 2003).

2. THE ELECTROENCEPHALOGRAM

The EEG is measured as the voltage difference between two or more electrodes on the surface of the scalp one of which is taken as a reference. The EEG expresses the overall activity of millions of neurons in the brain in terms of charge movement, but the electrodes can detect this only in the most superficial regions of the cerebral cortex. There are basically two conventions for positioning the electrodes on the scalp: the 10-20 Electrode Placement System (as recommended by the International Federation of Societies for EEG and Clinical Neurophysiology), and the Geodesic Sensor Net (developed by a firm called Electric Geodesics, Inc.). The former is more popular and is the convention adopted for the system described in this paper: it uses electrodes placed at positions that are measured at 10% and 20% of the head circumference (Figure 1b). In this case, the terminology for referring to the position of the electrodes uses a key letter that indicates a region on the scalp and a number: F = frontal, Fp = frontopolar, C = central, T = temporal, P = parietal, O = occipital and A = auricular (the ear lobe; not shown in Figure 1b). Odd numbers are for electrodes on the left side of the head and even numbers are for those on the right side. The set of electrodes being recorded at one time is called a montage. Montages normally fall into one of two categories: referential or bipolar. Referential means that the reference for each electrode is in common with other electrodes; for example, each electrode may be referenced to an electrode placed on the earlobe. An average reference means that each electrode is compared to the average potential of every electrode. Bipolar means that each channel is composed of two neighbouring electrodes; for example, channel 1 could be composed of Fp1-F3, where Fp1 is the active electrode and F3 is the reference, then channel 2 could be composed of Fp2-F4, where Fp2 is the active electrode and F4 is the reference; and so forth.

It takes many thousands of underlying neurons, activated together, to generate EEG signals that can be detected on the scalp. The amplitude of the EEG signal strongly depends on how synchronous is the activity of the underlying neurons. The EEG is a difficult signal to handle because it is filtered by the meninges (the membranes that separate the cortex from the skull), the skull and the scalp before it reaches the electrodes. Furthermore, the signals arriving at the electrodes are sums of signals arising from many possible sources, including artifacts like the heartbeat and eye blinks. Although medical specialists can diagnose brain malfunctioning from raw EEG plots, this signal needs to be further scrutinized with signal processing and analysis techniques in order to be of any use for a BCI system. There are a number of approaches to EEG analysis, such as power spectrum, spectral centroid, Hjorth, event-related potential (ERP), principal component analysis (PCI) and correlation, to cite but a few. Although powerful mathematical tools for analysing the EEG already exist, we still lack a good understanding of their analytical semantics in relation to musical cognition. However, continual progress in the field of Neuroscience of Music (Peretz and Zatorre 2003) is substantially improving this scenario. Once these issues are better understood we will be able to recognise patterns of cognitive activity in the brainwaves and activate appropriate musical algorithms associated with such patterns. Preliminary work in this regard has been reported in (Miranda et al. 2003;

Miranda et al. 2004). Brief non-mathematical introductions to EEG power spectrum and Hjorth analyses are given below due to their relevance to the system presented in this paper.

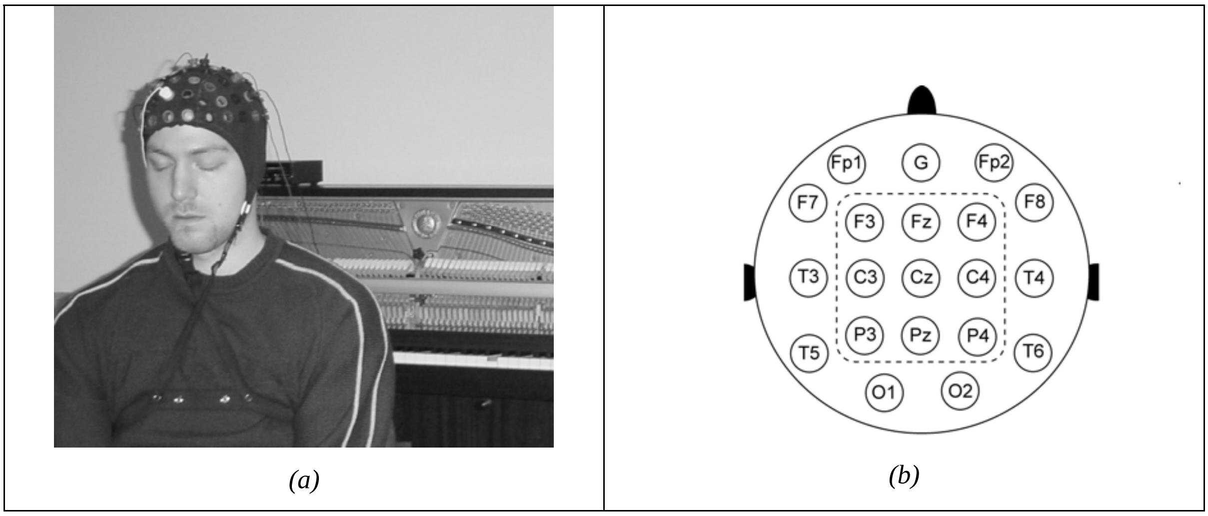


Figure 1. *(a) The BCMI system in action. Note the keys of the piano being played by the system; no hands or any kind of physical movement is needed. (b) The 10-20 electrode placement system.*

Table 1: *Electroencephalographic rhythms.*

Rhythm	Bandwidth	Meaning
Delta	Lower than 4 Hz	Sleep (non-dreaming)
Theta	Between 4 Hz and 7 Hz	Drowsiness
Alpha	Between 8 Hz – 13 Hz	Relaxed; aware but with eyes closed
Beta	Higher than 13 Hz	Awake, alertness, intense mental activity

Power spectrum analysis is derived from techniques of Fourier analysis, such as the Discrete Fourier Transform (DFT). In short, DFT analysis breaks the EEG signal into different frequency bands and reveals the distribution of power between them. This is useful because the distribution of power in the spectrum of the EEG reflects distinct “mental states”. In general, a signal characterised by low-frequency components (e.g., lower than 8 Hz) may be associated with a state of drowsiness, whereas a signal characterised by high-frequency components (higher than 14 Hz, up to 30 Hz or so) may be associated with a state of alertness. When the cortex is most actively engaged in processing information, whether generated by sensory input or by some internal process, the activity level of cortical neurons is relatively high, but also relatively unsynchronized. By contrast, when the cortex is less actively engaged in processing information, the activity level of the cortical neurons is relatively low, but also relatively synchronized. The higher the synchronization of the neurons, the higher the overall amplitude of the EEG (Giannitrapani 1985). There are four recognised frequency bands of EEG activity, also referred to as EEG rhythms, associated with different mental states in healthy adult brains, as shown in Table 1 (Misulis 1997).

Hjorth (1970) introduced an interesting method for clinical EEG analysis, which measures three attributes of the signal: its activity, mobility and complexity. Essentially, Hjorth analysis is a time-based amplitude analysis. This method is interesting because it represents each time step (or window) using only these three attributes and this is done without conventional frequency domain description. The signal is measured for successive epochs (or windows) of one to several seconds. Two of the attributes are obtained from the first and second time derivatives of the amplitude fluctuations in the signal. The first derivative is the rate of change of the signal’s amplitude. At peaks and troughs the first derivative is zero. At other points it will be positive or negative depending on whether the amplitude increases or decreases with time. The steeper the slope of the wave, the greater will be the amplitude of the first derivative. The second derivative is determined by taking the first derivative of the first derivative of the signal. Peaks and troughs in the first derivative, which correspond to points of greatest slope in the original signal, result in zero amplitude in the second derivative, and so forth. Activity is the variance of the amplitude fluctuations in the epoch. Mobility is calculated by taking the square root of the variance of the first derivative divided by the variance of the primary signal. Complexity is the ratio of the mobility of the first derivative of the signal to the mobility of the signal itself. There is no clear agreement as to what these measurements mean in terms of cognition. It is common sense to assume that the longer a subject remains focused on a specific mental state, the more stable the signal is, and therefore the variance of the amplitude fluctuation is lower.

3. THE GENERATIVE MUSIC ENGINE

Our system features a machine-learning algorithm that learns generative music rules from given examples (i.e., musical scores in machine-readable format). Due to limited availability of space, this paper focuses only on the functioning of the generative component of the system.

The EEG signal can influence in a well defined way the mixture of different style-elements found in the different musical examples given to train the system. It can generate music that contains, for example, more Schumann-like elements when the spectrum of the subject's EEG is characterised by alpha rhythms or more Beethoven-like elements when the spectrum of the EEG is characterised by beta rhythms; these associations are arbitrary.

Table 2: *An excerpt from a database of musical elements where: CO = composer (SCHU = Schumann.), P-CLASS = pitch class, P = pitch, PCL = pitch-class leading, PL = pitch leading and TYPE = type.*

ID	**SCHU-1-1-CAD**	**ID**	**SCHU-1-1-MEA-6**
CO	**SCHU**	**CO**	**SCHU**
P-CLASS	**((0 2 7) (0 2 4 5 7 11))**	**P-CLASS**	**((5 9) (0 5 7 9))**
P	**74**	**P**	**81**
PCL	**((0 4 9) (0 2 4 5 7 9 11))**	**PCL**	**((0 2 7) (0 2 4 5 7 11))**
PL	**76**	**PL**	**74**
TYPE	**CAD**	**TYPE**	**BAR**
ID	**SCHU-1-1-MEA-1**	**ID**	**SCHU-1-1-MEA-4**
CO	**SCHU**	**CO**	**SCHU**
P-CLASS	**((0 4) (0 3 4 6 7 9))**	**P-CLASS**	**((0 4) (0 3 4 6 7 9))**
P	**76**	**P**	**83**
PCL	**((2 7 11) (2 5 7 9 11))**	**PCL**	**((0 4) (0 4 7))**
PL	**83**	**PL**	**76**
TYPE	**INC**	**TYPE**	**BAR**
ID	**SCHU-1-1-MEA-3**	**ID**	**SCHU-1-1-MEA-2**
CO	**SCHU**	**CO**	**SCHU**
P-CLASS	**((0 4) (0 3 4 6 7 9))**	**P-CLASS**	**((2 7 11) (2 5 7 9 11))**
P	**76**	**P**	**83**
PCL	**((2 7 11) (2 5 7 9 11))**	**PCL**	**((0 4) (0 3 4 6 7 9))**
PL	**83**	**PL**	**76**
TYPE	**BAR**	**TYPE**	**BAR**

Example-based musical-generation systems are often based on formalisms such as transition networks or Markov Chains to re-create the transition-logic of what-follows-what, at the level of notes and at the level of similar "vertical slices" of music. The act of recombining the building blocks of music material together with some typical patterns and structural methods has proved to have great musical potential (Cope 2001). Self-learning predictors of musical elements based on previous musical elements can be used at any level or for any type of musical element such as: musical note, chord, bar, phrase, section, and so on. The current version of the system uses a simple statistical predictor at the level of short vertical slices of music such as a bar or half-bar, where the predictive characteristics are determined by the chord (harmonic set of pitches, or pitch-class) and by the first melodic note following the melodic notes in those vertical slices of music. The system uses a method for generating short musical phrases with a beginning and an end that allows for real-time steering with information extracted from the EEG. The system generates music by defining top-level structures of sequences and methods of generating similarity- or contrast-relationships between elements. Consider the following example (LISP-like notation):

S -> (INC BAR BAR BAR BAR BAR HALF-CADENCE 8BAR-COPY)

From this top-level the system generates rules for selecting a valid musical building block for each symbol, including rules for incorporating the EEG information in all decisions. For example:

```
INC ->     ( (EQUAL 'MEASURE-1) (EQUAL 'COMPOSER  EEG-SET-COMPOSER) )

BAR -> ( (CLOSE 'PITCH 'PREV-PITCH-LEADING)
         (CLOSE 'PITCH-CLASS 'PREV-PITCH-CLASS-LEADING)
         (EQUAL 'COMPOSER EEG-SET-COMPOSER) )
```

This defines a framework to generate a valid sequence with a beginning and an end, including real-time EEG control through the variable **EEG-SET-COMPOSER**. The generative engine will find a musical element for each of the constraint-sets that are generated above from **INC** and **BAR**, by applying the list of constraints in left-to-right order to the set of all musical elements until there are no constraints left, or there is only one musical element left. This means that some of the given constraints might not be applied.

The database of all musical elements contains information from different composers (or musical styles), with elements tagged by their musical function such as “measure 1” for the start of a phrase, “cadence” for the end, “composer” for the composer or musical style, and the special tags “pitch” and “pitch-class” that are both used for correct melodic and harmonic progression or direction. As an example, Table 2 lists excerpts from a database showing the main attributes that are used to recombine musical elements.

P-CLASS (for pitch-class) is a list of two elements. The first is the list of start-notes, transposed to the range of 0-11. The second is the list of all notes in this element (also transposed to 0-11). **P** is the pitch of the first (and highest) melodic note in this element. By matching this with the melodic note that the previous element was leading up to, we can generate a melodic flow that adheres in some way to the logic of “where the music wants to go”. The **PCL** (for pitch-class leading) elements contain the same information about the original next bar; this is used to find a possible next bar in the recombination process. Then there are the **INC**, **BAR**, and **CAD** elements. These are used for establishing whether those elements can be used for phrase-starts (incipient), or cadence. Simply by combining the musical elements with the constraint-based selection process that follows from the terminals of the phrase-structure rewrite-rules, we end up with a generative method that can take into account the EEG information. This generates musical sequences with domino game-like building block connectivity:

```
( (EQUAL 'MEASURE-1) (EQUAL 'COMPOSER EEG-SET-COMPOSER) )
```

Assuming that there are also musical elements available from composers other than **SCHU**, the first constraint will limit the options to all incipient measures from all musical elements from all composers. The second constraint will then limit the options according to the current EEG analysis to the composer that is associated with the current EEG activity, as follows:

```
( (CLOSE 'PITCH 'PREV-PITCH-LEADING)
  (CLOSE 'PITCH-CLASS 'PREV-PITCH-CLASS-LEADING)
  (EQUAL 'COMPOSER EEG-SET-COMPOSER) )
```

Figure 2. *An example of a generated mixture of Robert Schumann and Ludwig van Beethoven.*

In the given phrase structure, the rule that follows from **BAR** then defines the constraints put upon a valid continuation of the music. These constraints will limit the available options one by one, and will order them according to the defined rule preferences. The **CLOSE** constraint will order the available options according to their closeness to the stored value. For example, after choosing (refer to Table 2): **(SCHU-1-1-MEA-1 CO SCHU P-CLASS ((0 4) (0 3 4 6 7 9)) P 76 PCL ((2 7 11) (2 5 7 9 11)) PL 83 TYPE INC)** as the beginning, **PREV-PITCH-LEADING** will have stored 83, and **PREV-PITCH-CLASS-LEADING** will have stored **((2 7 11) (2 5 7 9 11))**. This will result in measures 2 **SCHU-1-1-MEA-2** and 4 **SCHU-1-1-MEA-4** being ranked highest according to both pitch and pitch-class, and measure 6 and the cadence **SCHU-1-1-CAD** close according to pitch-class, while measure 6 **SCHU-1-1-MEA-6** is also quite close according to pitch. This weighted choice will give a degree of freedom in the decision that is needed to generate pieces with an element of surprise. The music will not get stuck in repetitive loops when no perfect match is available, but it will find the closest possible continuation. The system can still find a close match this way if the third constraint eliminates all the obvious choices that are available; e.g., because a jump to the musical elements of another composer or style is requested, which might not use the same pitch-classes and pitches. Figure 2 shows an example output with elements from the musical styles of Robert Schumann and Ludwig van Beethoven. In this example, the EEG jumped back and forth from bar to bar between the two styles.

4. THE BRAIN-COMPUTER MUSIC INTERFACE SYSTEM

Our BCMI system falls into the category of BCI computer-oriented systems (Miranda et al. 2003). These systems rely on the capacity of the users to learn to control specific aspects of their EEG, affording them the ability to exert some control over events in their environments. Examples have been shown where subjects learn how to steer their EEG to select letters for writing words on the computer screen (Birbaumer et al. 1999). However, the motivation for our research departs from a slightly different angle to other BCI systems. We aimed for a system that would make music by “guessing” the meaning of the EEG of the subject rather than a system for explicit control of music by the subject. We acknowledge unreservedly that the notion of “guessing the meaning of the EEG” might be rather naïve. Nevertheless we took the risk of suggesting this notion because it is a plausible notion: it is based on the assumption that neurophysiologic information can be associated with specific mental activities (Anderson and Sijercic 1996; Petsche and Etlinger 1998). Continual progress in the field of Neuroscience of Music is increasingly substantiating this assumption (Peretz and Zatorre 2003).

The BCI research community understands that a BCI system is a system that allows for the control of a machine by explicitly imagining the task(s) in question; e.g., control a robotic arm by thinking explicitly about moving an arm. This is a very difficult problem. The system presented in this paper does not address this type of explicit control. Moreover, we are not interested in a system that plays a melody by imagining the melody itself. Rather, we are interested in furnishing our systems with Artificial Intelligence in order to allow them to make their own interpretation of the meaning of the EEG patterns. This is a fundamentally different approach to BCI because such machine-interpretations may not always be accurate or realistic. However, this is exactly the type of man-machine interaction that we are interested in exploring. The system is programmed to look for information in the EEG signal and match the findings with assigned generative musical processes corresponding to different musical styles. As mentioned in the previous section, these assignments are arbitrary.

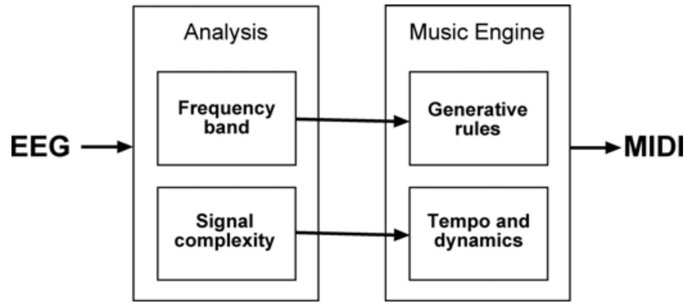


Figure 3. *Spectral information is used to activate generative music rules to compose music on the fly and the signal complexity is used to control the tempo and dynamics of the performance.*

The EEG is sensed with seven pairs of gold EEG electrodes on the scalp, forming a bipolar montage: F8-T4, F4-C4, Fz-Cz, F3-C3, F7-T3, P3-O1 and P4-O2. A discussion for the rationale of this configuration falls beyond the scope of this paper. It suffices to say that we are not looking for signals emanating from specific cortical sites; rather, the idea is to sense the EEG over a wide area of the cortex. The electrodes are plugged into a biosignal amplifier and a real-time acquisition system. The analysis module performs power spectrum and Hjorth analyses in real-time. It generates two streams of control parameters. One stream contains information about the most prominent frequency band in the signal and is used to generate the music. The other stream contains information about the complexity of the signal and is used to control the tempo (beat) and dynamics (loudness) of the performance (Figure 3).

The present implementation activates generative rules for two different styles of music (e.g., Schumann-like and Beethoven-like), depending on whether the EEG is characterised by the presence of alpha rhythms or beta rhythms. Every time it has to produce a musical bar (or half-bar), it checks the power spectrum of the EEG at that moment and activates the generative rules accordingly. The system is initialised with reference tempo and amplitude values, which are modulated by the signal complexity analysis (Hjorth analysis).

5. CONCLUDING REMARKS

In order to have greater explicit (or intentional) control over the system, we are studying ways to develop methods to train subjects to achieve specific EEG patterns to control musical algorithms. We have initial evidence that this can be made possible using a technique known as biofeedback. Biofeedback is when biological information or signals are monitored electronically, which in turn feeds back information about your body or physiological state. This information is often displayed through audio-visual stimuli. As a result, the subject can learn to modify these signals and subsequently learn to gain greater control of the biological signals. Biofeedback technology has been used to treat and control a number of conditions; examples include migraine headaches and epilepsy. However, the problem of explicit objective control "by thinking" is a difficult one because while EEG analysis can help us to deduce *whether* a person is thinking, it is unlikely to give us any clear indication of *what* a person is actually thinking.

We acknowledge that the music produced by the system is of limited appeal for those interested in modern new music. Furthermore, the pieces produced by our computer-replication of musical style engine may not always sound convincing to discerning listeners. However, we decided to adopt the grammar-based approach inspired by the work of David Cope (2001) as a starting point for this research because such grammars are well understood and their use in music is well documented. Nevertheless, we are studying the possibility of using other interesting machine learning and generative techniques such as those proposed by Assayag and Dubnov (2004).

An aspect that needs to be addressed in future versions of our system is the non-ergonomic nature of the electrode technology for sensing the EEG. It can be uncomfortable and awkward to wear. There are various possibilities for innovations in the hardware design of EEG capture devices. Inexpensive auto-scanning and auto-negotiating wireless chips are now available and could be placed on the head along with the small preamplifiers. It is thus technically possible to build wearable EEG amplifiers with built-in signal processing and wireless data transmission.

Acknowledgements: The author would like to thank two former members of his post-graduate research team, Andrew Brouse and Bram Boskamp, for their valuable contributions to this project.

6. REFERENCES

E D Adrian and B H C Matthews (1934), The Berger Rhythm: Potential Changes from the Occipital Lobes in Man, *Brain*, **57**, 4, pp. 355-85.

C Anderson and Z Sijercic (1996), Classification of EEG signals from four subjects during five mental tasks, *Solving Engineering Problems with Neural Networks: Proceedings of the Conference on Engineering Applications in Neural Networks (EANN'96)*, London, UK.

G Assayag and S Dubnov (2004), Using Factor Oracles for Machine Improvisation, *Soft Computing*, **8**, 9, pp. 604-610.

H Berger (1929), Über Das Elektrenkephalogramm Des Menschen, *Archiv für Psychiatrie und Nervenkrankheiten*, **87**, pp. 527-70.

N Birbaumer, N Ghanayin, T Hinterberger, I Iversen, B Kotchoubey, A Kubler, J Perelmeouter, E Taub and H Flor (1999), A spelling device for the paralysed, *Nature*, **398**, pp. 297-298.

D Cope (2001), *Virtual Music*. Cambridge, MA: The MIT Press.

D Giannitrapani (1985). *The Electrophysiology of Intellectual Functions*. Basel: Karger.

C Guger and G Pfurtscheller (1998), EEG based brain-computer interface (BCI) running in real-time under Windows, *Electroencephalography and Clinical Neurophysiology (Proceedings of 9th European Congress of Clinical Neurophysiology)* **106**, 1, pp. 91-100.

B Hjorth (1970), EEG analysis based on time series properties, *Electroencephalography and Clinical Neurophysiology*, **29**, pp. 306-310.

T Kohonen, P Laine, K Tiits and K Torkkola (1991), A Nonheuristic Automatic Composing Method, In *Music and Connectionism* (P Todd & D G Loy, Eds), Cambridge (MA): The MIT Press.

A Lucier (1976), Statement On: Music for Solo Performer, In *Biofeedback and the Arts, Results of Early Experiments* (D Rosenboom, Ed), Vancouver: Aesthetic Research Center of Canada Publications.

E R Miranda, S Roberts and M Stokes (2004), On Generating EEG for Controlling Musical Systems, *Biomedizinische Technik*, **49**, 1, pp. 75-76.

E R Miranda, K Sharman, K Kilborn and A Duncan (2003), 'On Harnessing the Electroencephalogram for the Musical Braincap', *Computer Music Journal*, **27**, 2, pp. 80-102.

K E Misulis (1997), *Essentials of Clinical Neurophysiology*. Boston (MA): Butterworth-Heinemann.

E Niedermeyer and F H Lopes da Silva (Eds.) (1987), *Electroencephalography*, 2nd edition. Munich: Urban and Schwartzenberg.

I Peretz and R J Zatorre (2003), *The Cognitive Neuroscience of Music*. Oxford: Oxford University Press.

H Petsche and S C Etlinger (1998), *EEG and Thinking*. Vienna: Austrian Academy of Sciences.

D Rosenboom (Ed.) (1975), *Biofeedback and the Arts, Results of Early Experiments*. Vancouver: Aesthetic Research Center of Canada Publications.

D Rosenboom (1990), The Performing Brain, *Computer Music Journal*, **14**, 1, pp. 48-65.

R Teitelbaum (1976), In Tune: Some Early Experiments in Biofeedback Music (1966-1974), In *Biofeedback and the Arts, Results of Early Experiments* (D Rosenboom, Ed), Vancouver: Aesthetic Research Center of Canada Publications.

J J Vidal (1973), Toward Direct Brain-Computer Communication, In *Annual Review of Biophysics and Bioengineering* (L J Mullins, Ed.), pp. 157-180.

J Wolpaw, D McFarland, G Neat, C Forneris (1991), An EEG-Based Brain-Computer Interface for Cursor Control, *Electroencephalography and Clinical Neurophysiology*, **78**, 3, pp. 252-9.

Investigating the use of force feedback joysticks for low cost robot-mediated therapy

H Sugarman[1], E Dayan[2], A Lauden[3], A Weisel-Eichler[4] and J Tiran[5]

[1,4]Ono Academic College,
104 Tzahal St., Kiryat Ono, ISRAEL

[2]Sonarion-Hadassah Academic Virtual Reality Center, Hadassah College Jerusalem,
P.O. Box 1114, Jerusalem, ISRAEL

[3]Dept. of Biomedical Engineering, [5]Dept. of Mechanical Engineering,
Ben-Gurion University of the Negev, P.O. Box 653, Beer Sheva, ISRAEL

hsugarman@ono.ac.il, [2]*udi@sonarion.com*, *arnon.lauden@gmail.com,*
[4]*aviva.we@gmail.com*, *tiran@bgu.ac.il*

ABSTRACT

We are developing a low-cost robotic system– the Jerusalem Telerehabilitation System – using a force feedback joystick and a standard broadband internet connection. In this study, the system was found to be user-friendly by therapists, older and younger normal subjects, and a post-stroke subject. Kinematic analysis of the joystick movements showed differences between the older and younger normal subjects and between the post-stroke subject and older normal subjects. These preliminary data indicate that our low-cost and straightforward system has the potential to provide useful kinematic information to the therapist in the clinic, thereby improving patient care.

1. INTRODUCTION

Survivors of brain injury or stroke can improve movement ability with intensive, supervised training (Butefisch et al, 1995), and improvement with training may continue for a long time after injury (Forster and Young, 1996). However, economic pressures on health care systems and distance from rehabilitation centers prevent many patients from receiving the therapy they need. Robot-mediated therapy, which increases the time available for repetitive movement training, and telerehabilitation, the delivery of medical rehabilitation services at a distance via the internet, have the potential to address these problems (Rosen, 1999; Krebs et al, 2004; Fasoli et al, 2004).

In recent years, many researchers have investigated the use of computerized mechanical devices to automate movement therapy for neurological conditions (Krebs et al, 2004; Fasoli et al, 2004; Reinkensmeyer et al, 2002; Coote and Stokes, 2005; Broeren et al, 2004; Burdea et al, 2000; Jadhav and Krovi, 2004). Many of these devices utilize haptic interfaces. Robotic therapy has been found to significantly improve the movement ability of the affected upper limb in stroke patients (Fasoli et al, 2004; Coote and Stokes, 2005). In addition to providing a platform for intensive, repetitive movement therapy, the use of robots and other computerized mechanical devices also provides the opportunity for detailed kinematic analysis of patient movement. Kinematic analysis gives us information that we would not otherwise have. For example, Sugarman et al. (2002) demonstrated the presence of irregularities (segmentation) in the movement of the ostensibly unaffected hand in post-stroke subjects. Hwang et al (2005) had similar results showing segmentation of cursor trajectories. Quantitative information about movement (timing, joint angular and segmental paths, velocities and accelerations) can be used for diagnosis, evaluation and monitoring of patient progress (Liebermann et al, 2006), and would, therefore, be a useful tool for therapists in a clinical setting.

Despite the important information that can be derived from kinematic analysis, use of this type of analysis is not a routine part of treatment in most therapeutic settings. The reason is that most of the computerized systems that have been used for such analysis are complicated and expensive and, therefore, not suitable for ordinary clinic or home use. Moreover, in order to apply them for telerehabilitation use,

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

special phone lines may be required (e.g. Piron et al, 2004). We have followed the lead of the inexpensive Java Therapy System described by Reinkensmeyer et al (2002) and are developing a low cost robotic system - The Jerusalem Telerehabilitation System (Sugarman et al, 2006) - for the treatment of stroke and traumatic brain damage. The system uses a commercially available force feedback joystick, which we have programmed to either assist or resist the patient's movements, an ordinary home PC, and an arm rest we designed and built ourselves. The arm rest is designed in such a way as to enable patients to move the joystick by means of relatively large movements of their shoulder and elbow joints. Use of the arm rest also eliminates the need to use a splint to help the patient grasp the joystick. A graphical representation of the client's movements is available via the "graph animator". The system monitors the status and progress of the patient, and various parameters of his movement abilities, such as movement time, as well as accuracy and trajectory, and prepares reports for the patient and the therapist.

The Jerusalem Telerehabilitation System can work in two modes, a stand alone mode and a cooperative mode. In the stand-alone mode, the patient works either in the rehabilitation clinic or independently at home; in either case, a file recording the exercise session will be stored locally and uploaded automatically to a central data base at the end of the session. In the cooperative mode, patient and therapist are online at the same time and the therapist can observe the patient's movements and can guide the patient's joystick if necessary. In the cooperative mode, the system will work via a standard broadband internet connection.

Compared to existing systems, our system will have the following advantages: 1) The system will be inexpensive and easy to use. 2) The system will have remote monitoring and guidance of the patient's joystick by the therapist in the clinic. 3) Our system will provide a detailed report of kinematic variables such as timing, path length, straightness of path, number of sub-movements and degree of segmentation (see Sugarman et al, 2002); this will give the therapist in the clinic a practical, easy to use tool for tracking patient status and progress. 4) Instead of using Java applets that download each time (as in the Java therapy system), in our system the game is downloaded once and then the client can use it at will without waiting each time; client programs and data from exercise sessions are stored locally on the client's computer and uploaded to a central server at a later date for monitoring by the therapist. 5) There will be a "smart" system which self-adapts to the patient's capabilities in real time by increasing or decreasing the difficulty of the exercise as needed, for instance by changing the size of the target. 6) Subjects who are unable to grasp the joystick will use the custom-made arm rest that will allow them to move the joystick by means of movements of their shoulder and elbow joints, instead of by the usual method of small wrist movements; this will allow even severely involved subjects to use the system. 7) We plan to develop a central, international database which, by gathering data on many patients over time, will provide the basis for "smart" therapy and will also facilitate coordinated multi-center research studies.

As a preliminary test of our system, we conducted a pilot usability trial with one subject after stroke as well as several normal subjects of two different age groups. We also asked two physical therapists to try out the system. The goal was to see if therapists and subjects were able to use the system, and to examine the nature of the data obtained from the trials. We were particularly interested in investigating whether a simple, low cost system based on an off-the-shelf joystick with a relatively low sampling rate of 50 Hz would provide useful kinematic information.

2. METHODS

2.1 The system

The Jerusalem Telerehabilitation System consists of a client system and a therapist system (Figure 1). Currently, both systems are located in the clinic, but our plans are to make the client system suitable for use at home, connecting to the clinic via a standard broadband connection. Each system consists of a standard PC computer, a custom made arm rest and a force-feedback joystick as the input device.

We use Java for the graphical user interface and for communication functions. The therapeutic exercise uses Open GL to enable 3D visuals and DirectX for the sound and the joystick. There is a communication module in C++ for communication between the client and the therapist. This communication allows the therapist to see the movement of the client's joystick and also to guide the client's joystick, both in real time.

2.2 The subjects

There was one subject, age 65, with left hemiparesis, 10 years after stroke. There were three normal subjects in the age range 20-30 (younger group) and three normal subjects in the age range 50-65 (older group). The subject after stroke used the system as a preliminary usability test only. This trial did not affect his usual rehabilitation therapy. In this paper, we are not reporting on using our system as a treatment option. Two physical therapists, specializing in neurological disorders, were trained to use the system.

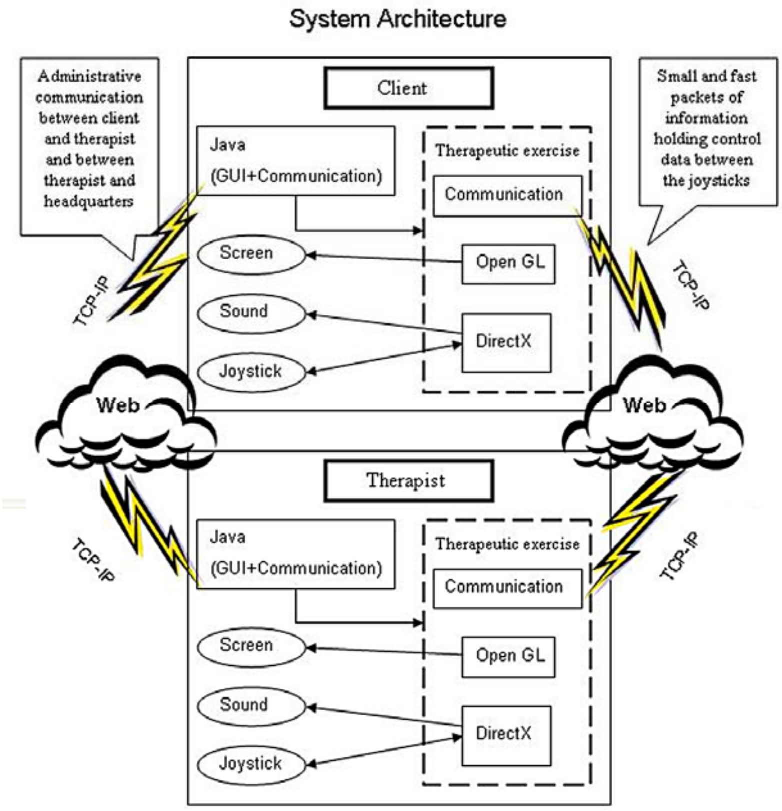


Fig 1. *Architecture of the JTR system. The client and the therapist each work with a PC and a joystick. They communicate with each other via the internet.*

2.3 The task

In the exercise that we tested, eight targets are arranged in a circle around the perimeter of the computer screen. (Figure 2) The subjects used the joystick to move the cursor to each of the 8 targets in turn, moving according to cue between opposite pairs of targets.

The motors in the force feedback joystick are programmed to either assist or resist patient movement. If the patients are unable to perform a movement, forces applied by the joystick will bring them very close to the target, and they then complete the movement on their own. If the patient is stronger, then the joystick is programmed to resist movement. The therapist can set the level of assistive or resistive force feedback as deemed appropriate for each patient. The maximum output of the joystick is approximately 5 newtons. In this study, subjects used the system both with and without force feedback. The kinematic analysis was done on those trials in which there were no forces applied to the joystick.

Each subject used the system for one session with three repetitions for each hand without force feedback. Afterwards, the subjects also tried out the various modes of force feedback – assistive and resistive at different strengths – and their reactions to the system were elicited.

2.4 Movement analysis

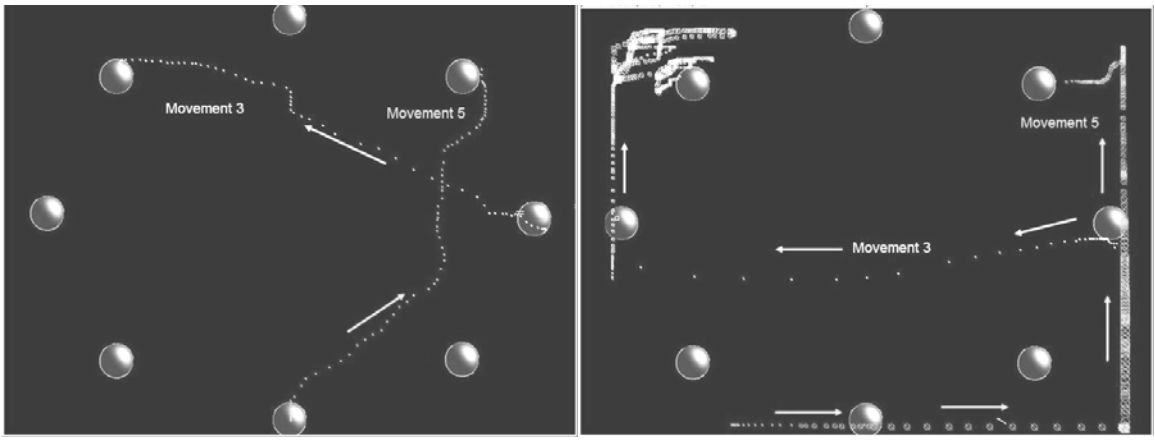
In each trial, the *x,y,* coordinates (in pixels) of the position of the joystick and the target position are sent by the joystick to the text file at a rate of 50 Hz. The forces applied to the joystick by the program are also sent to the text file at 50 Hz.

In this study, we examined the following kinematic parameters: movement time (average movement time between targets), and relative path length (the ratio of actual to ideal path length). In future studies, we plan to add other parameters such as analysis of the number of sub-movements in order to assess the degree of segmentation and thus the smoothness of the movements of brain damaged subjects.

3. RESULTS

3.1 Reactions of the subjects

Both normal subjects and the post-stroke subject reported that they enjoyed using the system. They were able to understand how to do the exercises and were able to concentrate on them. Subjects who were provided with feedback on the time it took them to complete the entire exercise (eight targets) said they enjoyed the competition aspect and tried to improve their time. The post-stroke subject asked when the system would be available for him to use at home. The two therapists who used the system found the program easy to learn and operate. Both initially expressed hesitation about using a computerized system; however, after a short (less than 1 hour) training session, both therapists were able to use the system independently. They had no difficulty remembering and using the various features of the program.



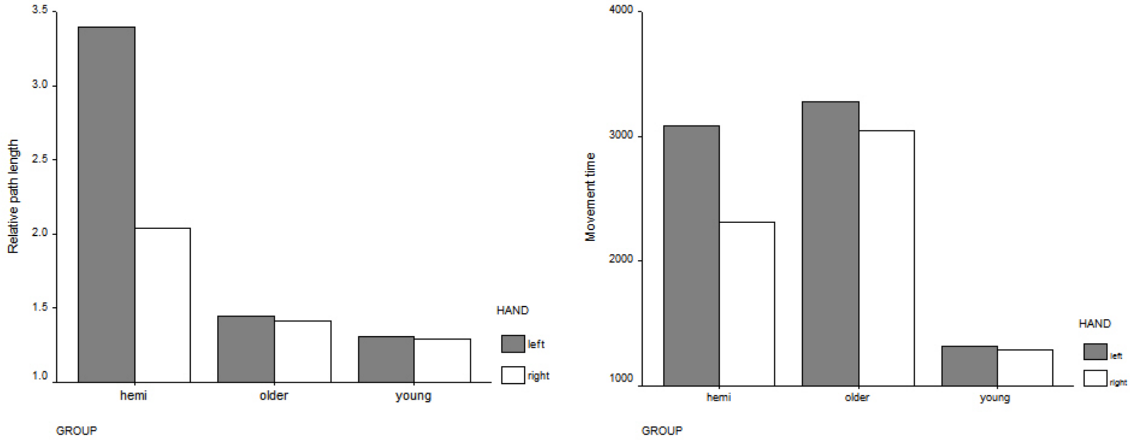
(a) *(b)*

Figure 2. *Screen shots from Graph Animator. The figure shows the eight targets arranged around the perimeter of the screen. These screen shots from the Graph Animator show movements 3 and 5 out of eight movements for a normal subject (a) and a post-stroke subject (b). Arrows indicate direction of movement. Distance between dots is proportional to the speed of the movement.*

3.2 Kinematic analysis

Kinematic analysis showed differences between the movements of the older and younger subjects and also indicated abnormalities in the movements of the post-stroke subject. The average relative path length for the older normal group was 1.4 (ratio of actual to ideal path length) for both hands, compared to 1.3 for the younger normal group (Fig. 3A). In addition, the average time for the individual movements to each target was longer for the older subjects (Fig. 3B). The post-stroke subject demonstrated much greater relative path length (3.4 and 2.0 for more affected and less affected hands, respectively) as compared to the normal older group. However, the average time for individual movements of the post-stroke patient was similar to that of normal subjects in his age group.

Examining the data from the post-stroke subject, we were able to differentiate between two types of difficulties: moving towards the target and homing in on the target. Figure 2A shows a normal subject performing movements 3 and 5 of the exercise and Figure 2B shows a record of a left hemiparetic subject performing the same two movements with his affected hand without the assistance of the joystick forces. It can be clearly seen that the post-stroke subject was not able to perform oblique movements and instead progressed by moving around the perimeter of the screen. In addition, we can see it was much easier for him to “home in” on the target on the top right as compared with the target on the top left.



(a) (b)

Fig. 3. *Kinematic parameters. A) The relative path lengths for the movement during the exercise was greater for the older normal subjects than for the younger group, and was much greater for the post-stroke subject. B) The average time for movement segments was greater for the older normal group than for the younger group; the average movement time for the post-stroke subject was similar to the older normal group.*

4. DISCUSSION

In this paper we have shown that the JTR system is capable of measuring differences in kinematic parameters such as movement time and relative path length. The trajectories of individual movements are recorded and can be analyzed for direction changes, difficulty in performing certain movements and the ability to home in on a particular target. This type of information − which movements and which directions are more difficult for each patient − will be useful to the therapist in customizing exercises for each patient. The system is simple and straightforward and was readily accepted both by therapists and the study subjects.

In recent years many studies have been done on kinematic abnormalities of movements following brain injury. Devices used include the Inmotion2 (commercial application of the MIT-Manus robot) (Rohrer et al, 2002), the Optotrak system (Michaelsen et al, 2006), the Elite 4 camera system (Van der Heide et al, 2005) and the Phantom (Broeren et al, 2002). Parameters evaluated included time, velocity, and relative path length (the degree of superfluous movements − trajectory/distance). These are all excellent systems; however, their high cost and the need for technical help in using them put them beyond the reach of the ordinary clinical setting. As far as we know, no low-cost easy to use systems are currently available. We would like to suggest that a possible spin-off of our simple, low-cost system would be to provide just this type of tool.

5. CONCLUSION

In conclusion, we found that this simple joystick paradigm can be used by patients and therapists and will give us useful kinematic information. The next step will be to do a wider clinical trial and develop more therapeutic exercises in preparation for distribution of the system in the homes of patients.

Acknowledgements. We are grateful to the Dorset Foundation for their generous funding of this project.

6. REFERENCES

J Broeren, M Rydmark and K S Sunnerhagen (2004), Virtual reality and haptics as a training device for movement rehabilitation after stroke: a single-case study, *Arch. Phys. Med. Rehabil.*, **85**, *8*, pp. 1247-1250.

G Burdea, V Popescu, V Hentz and K Colbert (2000), Virtual reality based orthopedic rehabilitation, *IEEE Trans. Rehabil. Eng.*, **8**, *3*, pp. 430-432.

C Butefisch, H Hummelsheim, P Denzler and K-H Mauritz (1995), Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand, *J. Neurolog. Sci.*, **130**, *1*, pp. 59-68.

S Coote and E K Stokes (2005), Effect of robot-mediated therapy on upper extremity dysfunction post-stroke – a single case study, *Physiother.*, **91**, pp. 250-256.

S E Fasoli, H I Krebs and N Hogan (2004), Robotic technology and stroke rehabilitation: translating research into practice, *Top. Stroke Rehabil.*, **11**, *4*, pp. 11-19.

A Forster and J Young (1996), Specialist nurse support for patients with stroke in the community: a randomised controlled trial, *Br. Med. J.*, **312**, *7047*, pp. 1642-1646.

C G Jadhav and V Krovi (2004), A low-cost framework for individualized interactive telerehabilitation. *Proc. 26th Annual International Conference in IEEE Engineering in Medicine and Biology Society (EMBS)*, San Francisco.

H I Krebs, M Ferraro, S P Buerger, M J Newbery, A Makiyama, M Sandmann, D Lynch, B T Volpe and N Hogan (2004), Rehabilitation robotics: pilot trial of a spatial extension for MIT-Manus, *J. Rehabil. Eng.*, **1**, *1*, pp. 5.

D G Liebermann, A S Buchman and I M Franks (2006), Enhancement of motor rehabilitation through the use of information technologies, *Clin. Biomech*. **21**, *1* pp. 8-20.

S M Michaelsen, R Dannenbaum and MF Levin (2006), Task-specific training with trunk restraint on arm recovery in stroke: randomized control trial. *Stroke*, **37**, *1*, pp. 186-192.

L Piron, P Tonin, E Trivello, L Battistini and M Dam (2004), Motor tele-rehabilitation in post-stroke patients. *Med. Inform.*, **29**, *2*, pp. 119-125.

D J Reinkensmeyer, C T Pang, J A Nessler and C C Painter (2002), Web based telerehabilitation for the upper extremity after stroke, *IEEE Trans. Neural Sys. Rehabil. Eng.*, **10**, *2*, pp. 102-108.

B Rohrer, S Fasoli, HI Krebs, R Hughes, B Volpe, WR Frontera, J Stein and N Hogan (2002), Movement smoothness changes during stroke recovery. *J. Neurosci.*, **22**, *18*, pp. 8297-8304.

M J Rosen (1999), Telerehabilitation, *NeuroRehabil.*, **12**, *1*, pp. 11-26.

H Sugarman, A Avni, R Nathan, A Weisel-Eichler and J Tiran (2002), Movement in the ipsilesional hand is segmented following unilateral brain damage, *Brain Cogn.*, **48**, *2-3*, pp. 579-587.

H Sugarman, E Dayan, A Weisel-Eichler and J Tiran (2006), The Jerusalem TeleRehabilitation System, a new, low-cost, haptic rehabilitation approach, *Cyberpsychol. Behav.*, **9**, *2*, pp. 178-182.

J C Van der Heide, J M Foch, B Otten, E Stremmelaar and M Hadders-Algra (2005), Kinematic characteristics of reaching movements in preterm children with cerebral palsy, *Pediatr. Res.*, **57**, pp. 883-889.

The design of a haptic exercise for post-stroke arm rehabilitation

E Lövqvist and U Dreifaldt

Interaction Design Centre (IDC), Computer Science and Information Systems
University of Limerick, IRELAND

erik.lovquist@ul.ie, *ulrika.dreifaldt@ul.ie*,

www.idc.ul.ie

ABSTRACT

In this paper we present an application based on an immersive workbench and a haptic device, designed to motivate stroke patients in their rehabilitation of their arm. The work presented in this paper is the result of a six-month project, based on evaluation with stroke patients and on informal interviews with medical doctors, physiotherapists and occupational therapists. Our application called "The Labyrinth" has been used for studies with stroke patients and we have seen that arm rehabilitation using Virtual Environments and haptics can be very encouraging and motivating. These factors are crucial to improve the rehabilitation process.

1. INTRODUCTION

For a stroke patient the main concern is to improve. Rehabilitation is often crucial for recovery and has to be performed on a daily basis (Carr and Shephard 2003). Every individual has different needs and preferences so it is important to have a wide range of exercises to choose from. One of the most common symptoms after a stroke is impaired arm movement and haptic devices can potentially be used in the rehabilitation process as an alternative to ordinary rehabilitation exercises.

According to research by Andersson et al. (2003) the patient normally makes good progress during the beginning of the rehabilitation program when there are regular meetings and practice sessions with the physiotherapist/ occupational therapist (PT/OT). The problems often begin when the patient is supposed to continue their rehabilitation at home and on their own. Lack of motivation due to boredom and lack of support from family and friends can make the rehabilitation progress slow down. According to Broeren (2002) the use of meaningful and rewarding activities has been shown to improve the patient's motivation to practice.

According to Andersson et al. (2003), a new system needs to be implemented to rehabilitate stroke. This is particularly urgent as a consequence of the reduction in hospital funding and the increasing queues for healthcare. They suggest moving more of the rehabilitation to the patient's home. To do this, the home rehabilitation exercises need to be introduced at the rehabilitation clinic, so that the patient can get used to and understand the exercises. The problem in this case is that rehabilitation at home is not as efficient as rehabilitation in a clinic working under the expertise of nursing staff, primarily because the rehabilitation at home is irregular and not undertaken frequently enough. This is mainly because patients find the exercises boring and not stimulating, and also because the feedback from the clinic is unsatisfactory and not frequent enough. If it was easier for patients to see their progress immediately then they would be more motivated to keep on doing their exercises.

According to Carr and Shephard (2003) computer games are likely to be increasingly used in training for various aspects of upper-limb movement. They state that the use of a game focuses attention on the outcome of the movement as opposed to the movement itself. The motivating effects of being an active participant on an interesting task may be powerful facilitators in the rehabilitative process.

During the work progress we visited a rehabilitation clinic in Sweden. We tried most of the current rehabilitation exercises ourselves. There were exercises such as baking a cake, building a jigsaw, planting seeds but also games like Othello as well as sports like dart, swimming and tennis. The overall feeling we got from performing the exercises was that they were boring, repetitive and gave us a feeling of being back at pre-school. We also talked to some of the patients and they had similar opinions. One patient said that she

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

felt the exercises sometime were only to tell her that she was worse than before the stroke, worse than average, and left her with a feeling of being worthless. Another patient said that she didn't do her best, because the exercises were not motivating enough. These two statements summarise the problem with the traditional exercises quite well. If the exercises leave the patients feeling worthless, it's quite obvious they will not feel motivated to do their best. If the patients are not doing their best, they will not receive the best possible rehabilitation.

The use of computer enhanced environments in combination with haptic devices for post-stroke rehabilitation has the possibility to address these problems and it is one of the reasons why it has become a popular area of research. Here we present some of them:

An advanced home rehabilitation systems incorporating both virtual reality and haptics (PHANTOM Omni) was proposed by Rydmark et al. (2005). In this system the patient's home-based workstation is connected to a rehabilitation unit with a game server with rehabilitation exercises, a patient database and a patient management system. Through the database system PT/OT can access information concerning motion capture data (movement of the haptic device). Using both data visualization and text output it is possible to analyze the work done by the patient while playing the game. The PT/OT can then evaluate the progress of the patient's rehabilitation and adjust the parameters of the game using the relevant patient data. System updates (game updates, new games, etc.) are carried out by the system developers when necessary. The benefit of this system is that new games may frequently be added and the PT/OT receives instant feedback on the patient's progress. It is used for rehabilitation of hemiplega (paralysis) in both the left and right upper-body. It also includes examples of games that are as yet very basic in nature. The same research team presented an initial study in 2002 (Broeren et al. (2002)) which shows promising results in the recovering process.

Another method used for post-stroke rehabilitation is "The Rutgers ankle" (Boian et al. (2002)).This involves controlling a computer game with a haptic device strapped to the foot. By moving the ankle, an airplane or boat is guided on-screen and the built-in-resistance helps rehabilitate the control mechanism of both the foot and the lower leg muscles. It is also worth noting that efforts have been made to create more engaging games by using force-feedback. For example, when the airplane bumps into something, as part of the game previously mentioned, the device gives the user haptic feedback.

An example of yet another device used for rehabilitation is the CyberGlove (Adamovich et al (2005)). It exercises various finger movements and increases strength. The exercises they use also seem to be quite simple of nature.

In this present project we have been trying to enhance our knowledge of the rehabilitative process through the exercises available today and creating a novel forms of exercise using virtual environments and haptics. It is not our intention to provide a substitute for those exercises already in use. We hope our application can complement them. The main reason for introducing virtual environments in combination with the rehabilitation process is the possibility of providing a number of alternative exercises that are engaging, stimulating and fun. In the research mentioned above, all exercises and games are quite basic and appear not being engaging enough. To fulfil the needs of a stroke patient, the exercises require careful design. In this paper we present an aid in the process of creating successful VR exercises in this context.

2. FUNDAMENTAL DESIGN CONSIDERATIONS

To make an exercise encouraging, stimulating, engaging and playable for a stroke patient several design aspects have to be considered. These aspects are the result of literature studies, informal interviews with medical doctors, physiotherapists, occupational therapists and a study with stroke patients. There are many design aspects and questions that arise when designing these exercises. In this paper we have decided to focus on the most important ones:

- *Reward system:* With an efficient reward system the chance of keeping the patients interest increases dramatically. Using a scoring system also gives the user instant feedback about the progress. By creating an exercise where the stroke patient can get a feel of achievement and encouragement, consequently the rehabilitation process will become more interesting. By incorporating the exercise in a VE it allows many possibilities of encouragement apart from just scoring systems, for example, a combined visual and sounding applause, cheerful animations and tunes, happy faces etc.
- *Difficulty:* The degree of disability after a stroke varies for different patients; hence the level of difficulty must be set individually. The patient's engagement needs to be maintained; too easy, no challenge; too hard, they are likely to give up. It is important to remember exercises that suit a smaller

range of patients should be considered as well. For example, specially designed exercises for patients in the first stages of the rehabilitation process where the patients may have particular needs.

- *Multimodal feedback:* It is important for the user to get real time task-specific feedback when interacting with the interface.
- *Environment design:* Textures and structures that resemble pleasant environments are effectively used to create suitable environments.
- *Intuitive task:* An exercise that would require long explanations and a long time of use to get used to might risk losing the patients interest in the initial stages. More advanced exercises could be motivated in the sense of increasing the challenge and make it more engaging, but must be designed carefully.
- *New possibilities:* In a virtual environment almost anything is possible. This allows us to listen to what stroke patients want and prefer, and create exercises based upon these criteria. There are many game ideas that stroke patients would find interesting, such as *The Labyrinth*, but also consider real-life situations that might be impossible to do with an impaired arm, for example playing an instrument, play sports, gardening etc. that are possible in a VE.

3. OUR APPLICATION

For our application we use an immersive workbench and a PHANTOM Omni from Sensable Technologies as a haptic input and output device to control the game. It is a mechanical arm that has 5 joints, enabling the user to interact with objects in the virtual environment. There is a ‘pen’ attached to the end of the mechanical arm and the user holds it with the hand to interact with virtual objects by moving their arm. The PHANTOM Omni is connected to a computer and the movement of the physical pen corresponds to the movement of the virtual stylus on screen (Figure 1).

As a development environment we chose to work with the H3D API from SenseGraphics AB. It is based on a scene graph structure and provides the developer with both haptic and graphical rendering. It runs on top of Open Haptics that is used for developing low level applications in C++. H3D is, on the other hand, a fairly easy and intuitive environment that enables a rapid development process. H3D API uses X3D for the scene graph and scripts are written in Python. The H3D API is open source and all updates are available freely on the Internet.

In the application we have built called “The Labyrinth” (Figure 1) the view for the user is as if the user was looking through a window into a room. Five outer walls make up the room and inside there are a number of smaller walls creating a maze.

The goal of the exercise is to navigate the virtual stylus through the maze (Figure 1), using the Phantom, from one panel to the other, without touching any of the walls. Touching the walls will give you minus five points but you are rewarded with plus fifty every time you touch the panels. After going back and forth two times a new set of walls are loaded, the next level. Every exercise consists of four levels, then you get a final score and a high score list based on previous games.

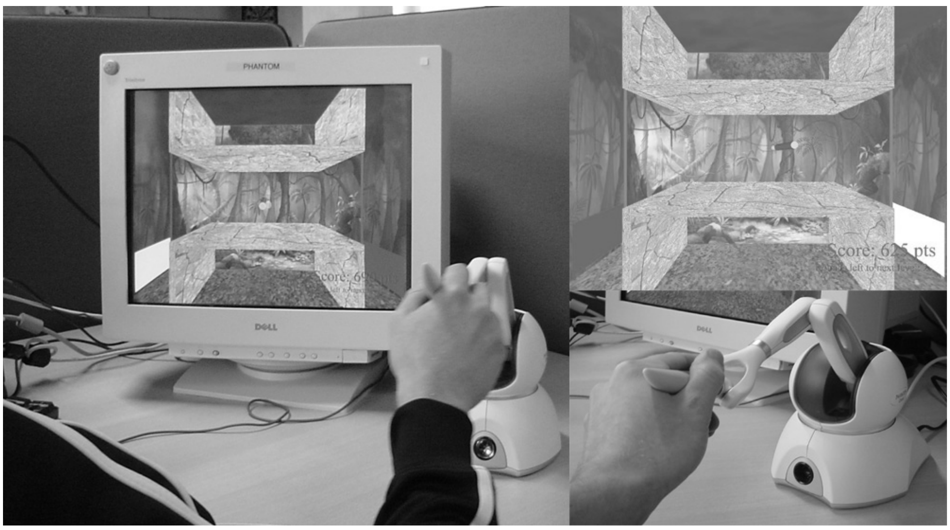


Figure 1. *The PHANTOM Omni and The Labyrinth.*

3.1 The design

The Labyrinth was created during an iterative design process and was informed by medical doctors, physiotherapists and occupational therapists. It was also finally tested with a group of stroke patients and after testing it was further refined. Here, we give some examples of how we have applied the fundamental design considerations in our application:

3.1.1 Reward System: The stroke patient gets rewarded and feedback on their progress by achieving a score that is displayed both during and at the end of a completed exercise. The top score from the five best games are displayed in a list. Ideally the patient feels motivated to set a higher score every time a new game is played and thereby wanting to improve. All of the scores are stored in a text-file making it possible for a PT/OT to monitor the patient's progress. This is the standard scoring method used in many computer games and works well for stroke exercises as well.

3.1.2 Difficulty: With The Labyrinth it is possible to set the difficulty by creating different mazes. It requires little effort and no programming to make new levels. This gives the PT/OT easy control over the level designs. In Figure 2 we present screen shots of four different levels of the game showing increasing difficulty.

It is the various movements e.g. avoiding the walls and going through the maze, that provides the training. A small maze with a few walls will make it possible for the user to go quickly through the maze using large arm movements, giving one type of training. Alternatively, the maze may have many walls and the user is forced to use smaller and more controlled movements thereby achieving another type of training.

Another way to set the degree of training is to use force fields. A force field is a static force applied in a specified direction and of a certain magnitude working on the haptic pen independent of its position. Initially, there is no force-feedback applied to the stylus, as it has no resistance. The task of going through the maze can then be made more difficult by applying resistance with a force field in, for example, the opposite direction of the patient's movement. Alternatively, a gravity force can be used to give the stylus a larger mass so it has to be lifted over the walls or balanced in between. There can also be objects that the user can pick up and carry to the other side of the labyrinth to gain bonus points. These objects can be of different size and weight to give a variety of training.

The Labyrinth is designed to be playable by a wide range of patients; both in the initial stages of rehabilitation and in their later stages.

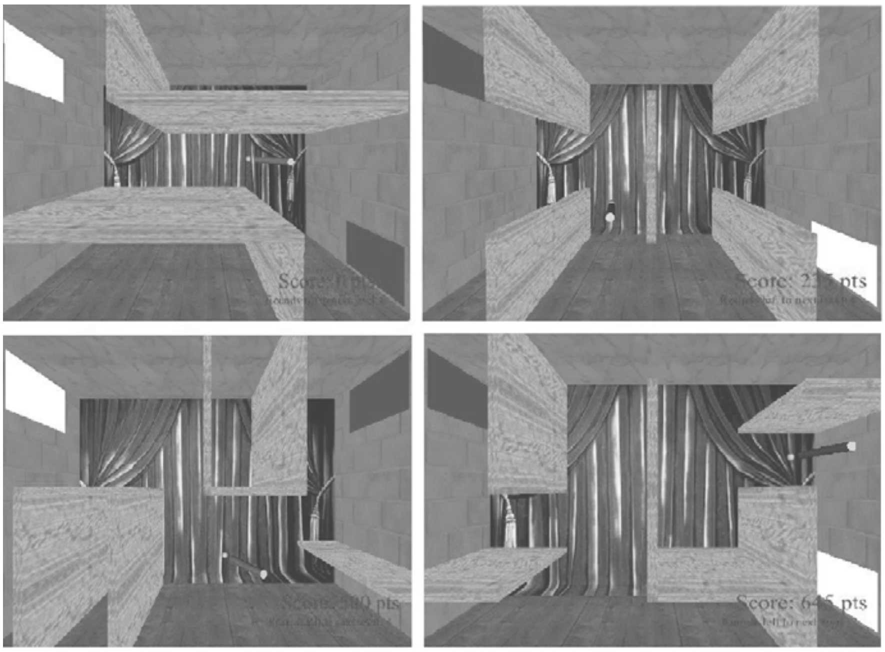


Figure 2. *Suggestion of level design with different difficulty.*

3.1.3 Multimodal Feedback. The PHANTOM Omni provides force feedback making it impossible to go through the walls. This forces the patient to stay inside of the maze and the user can feel when they are touching a wall. The user gets visual feedback every time one of the walls is touched by highlighting them in a different colour. Audio feedback is provided by using 'approving', harmonic sounds for the panels and 'disapproving', dissonant sounds for the walls. By using the force feedback as well as the audio and visual possibilities of the virtual environment, the communication with the stroke patient is very effective. The three different ways of showing the user when they have done something wrong makes the task more understandable. What we observed in our study (Dreifaldt and Lövquist 2006) was that patients want clear instructions before the exercise and instant feedback during the exercise, something that traditional exercises often lack.

3.1.4 Environment Design. In our case we have used three different types of environment to place the maze: jungle, theatre and mountains. Figure 2 shows the theatre design. Our goal is to find neutral but interesting environments.

3.1.5 Intuitive Task. With “The Labyrinth” the aim is to provide an exercise that is as intuitive as possible. It does not require more than a minute of explanation and then the patient can use it without hesitation.

4. STUDY

The initial prototype application design was based on literature reviews and informal interviews with medical doctors, physiotherapists and occupational therapists. To evaluate the application and to confirm our designs with end-users we did evaluation sessions with the “thinking aloud”-method (Lewis 1982) and additional discussions with three stroke patients. All of them enjoyed using “The Labyrinth” and they all showed increasing motivation when playing the game.

5.1 The Experimental setup

The final testing of the application was carried out during a one day qualitative study at a rehabilitation clinic in Sweden. The setup was the SenseGraphics 3D-IW Semi-immersive workbench and stereoscopical eyeglasses. The observations were both visual and auditory and the sessions were video recorded. The system was tested with three stroke patients. They all had 30 minutes each to use the system and another 30 minutes for an interview at the end of the session.

The users received the following instructions while watching the investigator playing the first level: Use the “Labyrinth” exercise. Start with level one, then move on to level two and finally level three. Start the game by pressing the “Start” button. The user should now move from one panel to the other and try not to touch any of the walls. The next panel to touch will light up (white) when touched and that panel then goes back to the “normal” colour and the other one lights up. Every time the users touch a panel they get plus 50 points and every time they touch a wall they get minus 5 points. When the level is finished the total score and a high-score list will be presented.

5.2 Users

The users are:

- *Patient A* was in her late 40’s and had a stroke two months ago. She had good movement in her arm but was still confined to a wheelchair. She had good computer knowledge and a positive attitude to computers even though she did think that computer games were mostly for children.
- *Patient B* was also in her late 40’s, she had her first stroke two years ago and was almost totally rehabilitated when she had a second stroke only a month before trying our system. She was the only one of the patients that was walking without a problem and also had the best movement in her arm. She didn’t have much computer knowledge and admitted she was hesitant of computers.
- *Patient C* was in his late 50’s and had a stroke only two weeks ago. Nevertheless he had a good movement in the arm. He told us he was very motivated to get better as soon as possible as he wanted to get back to his work as a surgeon. He was used to used to computers on a daily basis for writing patient’s journals, and he was open and interested to new technology in the medical field.

5.3 Observations

We did not experience any problems with their understanding of the stereoscopic display or how to use the PHANTOM; even the patient with the least experience of computers understood that the movement of the mechanical stylus corresponded to the movement of the virtual stylus on screen. After just a couple of minutes of explaining the PHANTOM and the game, they started playing without hesitation. All three were focused on the task and determined to do their best and achieve a good result. None of the patients had any problems understanding the concept of the game or seeing or understanding the virtual environment.

In our experiment we encouraged them throughout the application. All of the three patients were able to complete the exercise, but not too easily. They all agreed that the application had a good mixture of challenge and provided a feeling of accomplishment. In fact, they all said that it was the best exercise they had tried so far.

Even if all three patients seemed to enjoy the system they had different ways of assessing the system. One of them liked it because it felt like a meaningful exercise, another because it was a fun game. It does not seem to matter if the patient considers it a game, exercise or test as long as they engage with it and find it meaningful (Meyer (2001), Kielhofner (1995), Kielhofner (1997)). What we learned was that none of the patients wanted to give up, while two of them complained of sore muscles, but still kept on trying without complaining about the exercise.

Patient C said that he wanted to take the system home to continue training. He saw lots of potential in using it for training his finer motor skills. To him, the encouragement and the playfulness of the system was not as important as the actual training itself. The reward for him was seeing that using the PHANTOM could help him in his recovery. He also said that the points were not that important to him, the important task for him was to avoid the walls and not, as we had thought, to get as a high score as possible. For him it would have been better to get more feedback and encouragement during the game play.

They all found the idea of timing the exercises as something that could be used to increase the level of difficulty. This would be best done in the later stages of the rehabilitation when the patient is used to the exercise, making it more challenging and keep their interest. It was also important to have the possibility of choosing between playing with or without the timing facility. Generally we saw that the more encouragement we used, the better. This would suggest further system implementations incorporating more positive feedback during the game play. For example one could have a voice saying “Good work!” every time a level is completed or adding more possibilities for increasing the score.

The study was carried out in Swedish and below are some quotes freely translated by the investigator.

- *This is much more fun then I thought, I was a bit sceptical before I came here, computer games are mostly for children, but this is fun!* − Patient A.
- *It’s very fun to do this! I think this is the future, will be a great interest for this! If you look at the other inpatients here at the clinic there are of all age groups but a lot of young people. If you look into the future it will be even more young people and even more people used to use computers.* − Patient A.
- *You want it to be easy so you can handle it and gain some self-confidence, but not too easy, then it gets childish. But this is not too easy.* − Patient B.
- *With this one you get it confirmed right away that if you do something wrong, then you feel it and see it, very good!* − Patient B.
- *I think this was very fun; I like to have a system like this at home.* − Patient C.
- *I’m very enthusiastic; actually, this is a perfect tool for therapy.* − Patient C.
- *This is the most fun rehabilitation thing I’ve tried and moreover you practice fine motor skills and I’m dependent on that in my work, so it makes me very motivated.* − Patient C.

5. CONCLUSIONS AND FUTURE WORK

From our studies with stroke patients we have seen that our system contributes to an overall pleasant and encouraging exercise experience. Our main concern is that patients see this as a complement to the rehabilitation techniques used today and to give them alternative exercises that are encouraging, challenging and fun to help ease their recovery. We have shown that it is possible to create encouraging stroke rehabilitation exercises using virtual environments and we also produced a set of general guidelines that can be used in future designs.

The Labyrinth is an example of a haptic exercise for post-stroke arm rehabilitation. It is not designed to be “the perfect” exercise but an application to show many different design aspects that has to be taken into account. We believe that future work will be creating more exercises giving patients more choice. All patients are individuals and have different needs and preferences, hence there must be a wide range of exercises available.

The current application only uses force fields to make the game play more difficult. It would be desirable to create force fields to aid patients with severe arm disability to make it easier to perform the exercise. By having a force field that the stylus floats upon and repelling walls would make it easier to complete the task.

Acknowledgements: The project was approved by the Local ethical committee Göteborgs University, Sweden and the Research Ethics Committee at the University of Limerick, Ireland. We would also like to gratefully acknowledge the assistance of Mr. Paul Gallagher, Dr. Annette Aboulafia and Mr. Mikael Fernström, Interaction Design Centre, University of Limerick, Ireland.

6. REFERENCES

Adamovich S (2005), A Virtual Reality-Based Exercise System for Hand Rehabilitation Post StrokePresence. *Vol. 14, No. 2, April 2005*, 161-174.

Andersson T, Fredlund J, Hastad E, Härensten D, Karlsson J, Leterius M, Lundberg H, Sreand T-H (2003), Rehabliteringsstation för strokepatienter. Virtuella världar och haptika gränssnitt för rehabilitering i hemmet. Projekt *Produktutveckling, Produktion och produktionsutveckling*, Chalmers Tekniska Högskola tillsammans med Högskolan för Design och Konstverk, Göteborg 2003.

Boian et al. (2003), Haptic Effects for Virtual Reality-based Post-Stroke Rehabilitation., *Proceedings of the11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (HAPTICS'03)

Broeren J, Björkdahl A, Pascher R, RydmarkM (2002), Virtual Reality and Haptics as an Assessment Device in the Postacute Phase after Stroke . *CyberPshycology and Behavior*, Vol. 5, NO. 3, 2002

Carr J.H., Shepherd R. B. (2003), *Stroke rehabilitation: guidelines for exercise and training to optimize motor skill,* Butterworth-Heinemann, Oxford

Dreifaldt U., Lövquist E. (2006) The construction of a Haptic application in a Virtual Environment as a post-Stroke arm Rehabilitation exercise. *MSc thesis*, University of Linköping, Sweden.

Kielhofner, G. (1995) A Model of Human Occupation. Theory and application.. 2nd Ed. William & Wilkind, Baltimore.

Kielhofner, G. (1997) Conceptual foundations of occupational Theory. 2nd Ed. Davis, F.A. Philadelphia

Lewis, C. (1982). Using the "Thinking-aloud" Method in Cognitive Interface Design (Research Report No. RC 9265). Yorktown Heights: IBM.

Mayer A.(2001), The philosophy of occupational theraphy. *Am J Occup Ther*. 1922/1977;31:639-642. Am J Ouccup. Ther 2001;55:249-59

Rydmark M, Broeren J, Dixon M, Stibrant Sunnerhagen K (2005), Haptic Immersive Workbench for Stroke Rehabilitation.

SensAble Technologies, The USA, http://www.sensable.com/

SenseGraphics AB, Sweden, http://www.sensegraphics.se/

Investigating the impact of method of immersion on the naturalness of balance and reach activities

I Sander[1], D J Roberts[1], C Smith[2], O Otto[1] and R Wolff[1]

[1]Centre for Virtual Environments, [2]Centre for Rehabilitation and Human Performance Research, University of Salford, UK

irina.sander@web.de, {d.j.roberts, c.smith1, o.otto, r.wolff}@salford.ac.uk

[1]*www.nicve.salford.ac.uk*, [2]*www.healthcare.salford.ac.uk*

ABSTRACT

Immersive virtual reality is gaining acceptance as a tool for rehabilitation intervention as it places a person within a safe and easily configurable synthetic environment, allowing them to explore and interact within it through natural movement. The purpose of the study was to explore the usefulness of different types of virtual environments in the rehabilitation of upper limb function and balance in stroke patients. Although the above characteristics are ideal for rehabilitation of motor disorders, acceptance is hampered by insufficient knowledge of the effect of method of immersion on the naturalness of human movement. This study begins to address this by comparing the impact of two typical methods, Head Mounted Display (HMD) and immersive projection technology (IPT), on the naturalness of reach and balance activities. The former places the simulated image in front of the eyes, whereas the latter projects it around the user so that they perceive a holographic effect when wearing stereo glasses. Using the novel approach of placing the HMD in the IPT allowed subjects perceiving the environment through either, to be observed moving within the IPT holograph. Combined with sharing the same tracking and camera systems, this provided a direct comparison of tracking measurements, interaction behaviour, video and other observational data. The experiment studied subjects moving objects around a simulated living room setting initially on a level surface and then whilst varying the height and shape of the walking surface through raised planks. Performance in the synthetic environment, using both display types, was compared to that in a physical mock up of the living room. The experimental results demonstrate similar balance and reach movements in the physical mock up and the IPT, whereas a striking reduction in naturalness in both activities was observed for HMD users. This suggests that an inappropriate choice of method has the potential to teach unnatural motor skills if used in rehabilitation. Reasons for the difference are discussed along with possible remedies and considerations for practical applications within a clinical setting.

1. INTRODUCTION

Although immersive virtual reality technology has demonstrated substantial benefits in many sectors (Hughes, 2004), its potential has not been fully realised in many. The technology has been demonstrated particular attention in a range of both sensorimotor and cognitive training scenarios. Immersive virtual reality offers considerable potential for practice and assessment within rehabilitation, as it places a person within a safe and easily configurable synthetic environment allowing them to explore and interact within it through natural movement. To be effective in these domains, the technology must not alter the way in which people use their bodies. Many experts seem to hold strong yet differing opinions on the relative strengths and weaknesses of various methods of immersion but little if any research has directly compared the impact on naturalness of user movement. However, another concern in rehabilitation is the practicality of placement and maintenance of equipment within a clinical setting. Replacing a real environment with a virtual counterpart requires much larger and generally less robust technology, if the body is to be visible within it. For virtual reality to be widely adopted for rehabilitation we need a better understanding of the display factors that affect naturalness of body movement and to apply this in the design of systems for clinical settings. This paper starts to address the former and discuss the latter.

By placing a person within synthetic space in which body movement can be accurately monitored with respect to the simulated environment, immersive technologies offer particular potential in rehabilitation of

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3

patients with balance and reach disorders. Rehabilitation often requires individuals to relearn movement skills. Optimal ingredients for relearning have been described to be: “1) optimal sensory information 2) variability of practice 3) similarity between the context of training and the context of application” (Mulder & Hochstenbach, 2003). It is in the areas encompassed under 2) & 3) that virtual environments have the greatest contribution.

This study undertook a preliminary investigation of the impact of method of immersion on the naturalness of balance and reach activities. The approach was novel in that we observed subjects within a CAVE-like cubic Immersive Projection Technology (IPT) display, who either saw their body within the projected environment, or only the environment and some virtual representation of their hands and feet through a Head Mounted Display (HMD). This allowed straight forward analysis of body movement with respect to the environment and task. The experiments were carried out exclusively on subjects not suffering from motor disorders in order to assess risk and to understand the simple case of impact on able bodied people before adding the complex variables of disability and treatment.

The qualities of a VR system are typically measured in terms of Immersion and the user subjective illusion of Presence (Slater et al., 2001). The former relates to objective qualities such as the number of senses immersed in the simulation, the extent to which they are immersed, for example field of view, and the fidelity of representation, for example resolution. In contrast the latter describes a sense of being in the virtual, as opposed to the real environment (Meehan et al., 2002). Presence has many attributes and methods of measurement and the reader is pointed to the compendium of Presence Research (2005). This paper concentrates on the effect of immersion on the practice of the skill that is being learned. Therefore we focus on the faithfulness of body movement in response to given stimuli. Postural responses have previously been used in presence research as a behavioural objective corroborative measure of presence (Freeman et al., 2000). Here we observed the posture and reaching gestures of the subjects both at the time of the experiment and subsequently from video footage and tracked data.

The purpose of the study in this paper was to investigate the relationships between method of immersion, reach and balance. As this study concentrates on core issues that are relevant to a wider set of applications, we do not describe the rehabilitation system further in this paper. Related work of virtual reality in rehabilitation is surveyed with a focus on balance disorder and loss of motor control. A technical specification concisely describes the equipment used in the experiment. The experiment itself is then described, followed by results, discussion and conclusion.

1.1 Related Work

The comparison of IPT and HMD has not been widely studied. A lack of user studies is reported by Manek 2004 and a complete absence of direct comparisons reported by Steed 2005. The studies that we have found are restricted to selection, manipulation and locomotion. Reach and grasp is one method of selection. Once grasped, the object can be naturally manipulated. Locomotion refers to exploratory movement within the space, for example walking or turning. Manek found no previous literature on user studies comparing selection and manipulation techniques between the two methods of immersion (Manek, 2004) and showed that: selection and manipulation tasks can be affected by display type; that task performance can suffer when selection and manipulation techniques are migrated from one display type to the other; and that migrated techniques can be modified to compensate. A more recent study (Steed & Parker, 2005) reported that interaction techniques have been designed almost exclusively for HMDs, however, demonstrated that selection tasks were performed better in IPTs than in HMDs while little difference was noticed in that of manipulation tasks. The above comparisons were primarily focussed on the efficiency of object interaction; however, they did consider naturalness as an impacting factor. A correspondence between visual and proprioceptive senses is commonly seen as a major impacting factor on naturalness. Mine et al. (1997) presents a unified framework for VE interaction based on proprioception. Natural walking and turning is supported within a confined space by both display types, and the limits of the space can be overcome through manipulation of viewpoint through a hand held device. The typical configuration for an IPT is to have projection surfaces on three walls and a floor. Adding the fourth wall and roof significantly increase the complexity of the display and the space required to house it. One study showed that users of a typical IPT were less likely to use their body as opposed to the hand controller to turn (Bowman et al., 2002) and suggests that this is due to the missing wall. However, the configuration of the IPT was found to impact on orientation, moving and acting when comparing a six sided cubic display to a panorama (Kjeldskov, 2001). Balance disorders are disturbances that cause an individual to feel unsteady, giddy, woozy, or have a sensation of movement, spinning, or floating (NIDCD, 2005). Virtual Reality has been widely used for the treatment and rehabilitation for patients with the vestibular type of balance disorders. In such applications the aim is to produce a postural adjustment response in the viewer (Owen et al., 1998). Subjects can be observed to sway to compensate for moving images. This was demonstrated for stationary subjects in an HMD (Kuno et al., 1999) and both stationary and linearly walking subjects within an IPT (Keshner & Kenyon, 2000).

2. SPECIFICATION OF THE TEST ENVIRONMENT

A VR system uses real-time interactive graphics with three-dimensional models, combined with display technologies that give the user immersion in the model world and allow direct manipulation (Bishop & Fuchs, 1992). VR interfaces can include different displays systems, haptic interfaces, and real time motion tracking devices which are used to create environments allowing a user to interact with virtual objects in real-time through multiple sensory modalities. With an immersive environment, the software ensures that the visual scene's projection is always appropriate to the user's head positions to create the correct perspective view of three-dimensional objects and environment.

Immersion measures the extent to which sensual stimuli from the real world is replaced by synthetic stimuli. It is a multidimensional measure and considers the number of senses involved, the extent to which the real world is replaced and the fidelity of the synthetic stimuli (Slater et al., 2001). In order to maximise the accuracy of analyses or the transferability of skills learned in a VR rehabilitation tool, movement response to stimuli must be faithful to that in the real world. The previous section described other studies that found alignment between visual and proprioceptive senses to be an important factor in the naturalness of reach operations and it is reasonable to assume that it may also impact on balance. We therefore restrict our study to methods of immersion that attempt to align these two senses by surrounding the user in the displayed image. The two most common display types with this characteristic are the HMD and cubic IPT (e.g. CAVE). HMDs place displays in front of the eyes that move with the head. In contrast the IPT projects images onto large screens that surround the subject and are viewed through stereo glasses synchronised to a flicking offset in the image. Our approach to the comparison is novel in that we always observed the subjects within the IPT, even when they viewed the world through the HMD, allowing us a direct comparison of all recorded and observational data.

The two display types differ in a number of factors, most noticeably the subject only see his own body in the IPT. There are several additional attributes of HMDs that may have an effect on a user's performance. They come in a wide range of resolutions and different field of view (FOV). A lower FOV results in "tunnel vision" and might decrease immersion, but higher FOVs require spreading out the available pixels, which can decrease resolution and introduce distortion. In addition, there are ergonomic issues related to HMDs such as display size, weight and the ability to adjust various visual parameters (Bowman et al., 2002). Both of the displays used are better described as typical rather than state of the art. The IPT had four active surfaces, three walls and a floor and the HMD had a field of view of 60 degrees. In this case the IPT was more considerably more expensive the HMD. A study using state of the art rather than typical equipment would have used at least a five sided IPT and a HMD with at least a 110 degree field of view. Such luxuries were not available to us. A V8 HMD was used which has a resolution of 640x480. In contrast the IPT has a much wider field of view provided the subject does not turn towards an open wall. The previous section described how this characteristic can result in unnatural control of turning. Both systems use a tracking device to calculate the correct viewpoint and from this the stereo perspective. The primary hand is also typically tracked. Within our experiment we tracked the head, both hands and both feet. In order to draw direct comparison we used the same tracking system with both displays. The tracking system used was a magnetic Ascension Technologies Flock of Birds. We found that the registration (working volume) of the tracking system was insufficient to accurately monitor the feet from the default mounting within the IPT. We therefore placed tracking points below the knees and calculated foot position from these. Grasp-like selection of objects was achieved through a Pinch Glove on each hand capable of detecting a pinch between finger and thumb.

Within the IPT the user can see his own body, however representing a virtual hand that depicts grasp status can help to see what the system thinks is happening and thus cope with small tracking inaccuracies and temporary pinch glove failures. The HMD user can not see his own body and thus we are left with the choice of either giving no visual feedback of body position or of tying a virtual body to his movements. The latter is known to induce difficulties in accomplishing some tasks (Burns et al., 2005), however, the former can cause worse problems unless a close alignment between visual and proprioceptive feedback can be guaranteed. Given the complexities of ensuring this, including the accurate modelling and tracking of every limb, we decided to limit embodiment to hands and feet.

The simulation for the IPT ran on an SGI Onyx under the IRIX 6.5 operating system while that for the HMD ran on a Xenon PC with a dual-headed Nvida Quadro graphics card with Suse Linux 9.2. Apart from platform and display dependent components, the simulation and model were identical. The program was written in C/C++ and interfaces with the OpenGL Performer™ library version 3.0 scene graph library. The OpenGL Performer™ library supports multiple CPUs and provides a high-level graphics application programming interface (API). Tracking data and video were recoded with respect to the virtual environment throughout the experiment. The former recorded head, hand and feet movement. The video camera was placed just above head height at the entrance to the display.

3. EXPERIMENTATION

The purpose of the experiment was to investigate the impact of method of immersion on the naturalness of reach and balance. The scope was partly set by ethics and management of risk. We did not want to experiment on subjects with balance and reach disorders before we had assessed the impact of the technology on healthy subjects. This paper describes this initial experiment only. Furthermore, removing the variables of method and stage of treatment clarified the results.

3.1 Test Environment

A major challenge for rehabilitation is identifying effective and motivating intervention tools that enable transfer of the skills and abilities achieved within a VE to function in the real world. The importance of this in transferring sensorimotor skills has been demonstrated (Rose et al., 1996). For this reason we chose to simulate an everyday setting to which patients could relate and for which the transfer of skills would improve quality of life. The chosen simulated setting was a living room in which familiar objects could be picked up and moved around and various shaped planks placed above the floor for balancing. A basic physical mockup was built for comparison.

Four Scenarios were examined in the synthetic environment while varying the method of immersion. : repositioning of objects through reach, grasp, carry and placement; walking along a plank 2m over the floor; walking along a brittle plank 4m over the floor; and repositioning of objects while balancing on a 0.5m high curved plank. One and two handed object interaction were compared by simulating gravity in a way that required larger objects to be grasped by two hands before they could be moved. The first two scenarios were repeated with the physical mockup but for safety reasons the plank was only raised to 50cm above the floor.

The next subsections describe the particular experiments in more detail underlined with respective photographs taking during the trials.

3.2 Subjects

Twelve adults without any sensorimotor disorder age between 22 and 32, with different professional background and VR-knowledge have taken part in these experiments. Subjects were given time to familiarise themselves with the interface and basic activities and practiced each scenario before it was measured. This was necessary as people experiencing the VE for the first time had problems with judging the distance and functionalities of grasping. However, after a few training sessions the performance was considerable increased to close to that observed in the real world.

3.3 Tasks in Detail

3.3.1 Repositioning of objects through reach, grasp, carry and placement. In the first task subjects were asked to move various objects between predefined places in the room. With their right hand, subjects moved a potted plant, key board and small book, then moved the small book again with the left hand and finally moved a large book and then a chair with both hands. Before this first experiment, participants practised the handling of grasping virtual objects with the gloves, by recognising different states of the virtual hands and to interact with virtual objects. Afterwards, they were asked to perform a series of test-runs in a defined order.

3.3.2 Walking along a plank 2m over the floor. The second experiment measured balance but not reach and carry activities. The plank was rendered such that it appeared on the physical floor, whereas the floor of the room was rendered so that it appeared two metres below this. If the subject walked of the side of the plank, both plank and floor would raise so that the physical and virtual floor coincided. The task set was to walk up and down the plank several times without falling off.

The subjects were asked to start the balance experiments in there own time, thereby allowing them to familiarise and position themselves within the environment. This gave the participants the opportunity to feel more confident with the surrounding environment, and also avoided falling down before the experiment started.

Figure 1. *Balancing on a brittle plank 4m above the floor while observing the environment through the IPT.*

Figure 2. *Balancing on a shaped plank while observing the environment through the HMD.*

3.3.3 Walking along a brittle plank 4m over the floor. To measure the impact of increased anxiety, the above experiment was repeated for a much higher plank that would break if the subjects moved too quickly (Figure 1).

3.3.4 Repositioning of objects while balancing on a 0.5m high curved plank. Balance and carrying were measured together in the final test which was also made harder by using a curved plank (Figure 2). This time, only right hand manipulation was asked for.

4. RESULTS

Results presented here compare the naturalness of movement of subjects viewing the environment through HMD and IPT. The measurements presented in this paper are taken from tracking and video data. Tracking data describes the movement of feet and hands but not all experiments have logged both. Snapshots of the video data are presented in the paper and samples of the videos will be given in the presentation and have been submitted along with this paper. In terms of both tracked and video captured movements typical characteristics emerged across all test subjects that clearly differentiated the impact of method of immersion. In all experiments the typical visually observed and videoed movements of the IPT subjects indistinguishable from those in the physical mockup, while HMD subjects consistently demonstrated changed movement behaviour. The following subsections each describe typical samples.

4.1 Repositioning of objects through reach, grasp, carry and placement

Comparing tracked movements while repositioning objects, the curve of hand trajectories appears to be slightly smoother and considerably less erratic for the IPT users. This is most noticeable when looking from above at objects is being carried across the room, Figure 3. The videos are more telling as they clearly demonstrate exaggerated bends in the legs and back of subjects wearing the HMD, Figure 6.

4.2 Walking along a plank 2m over the floor

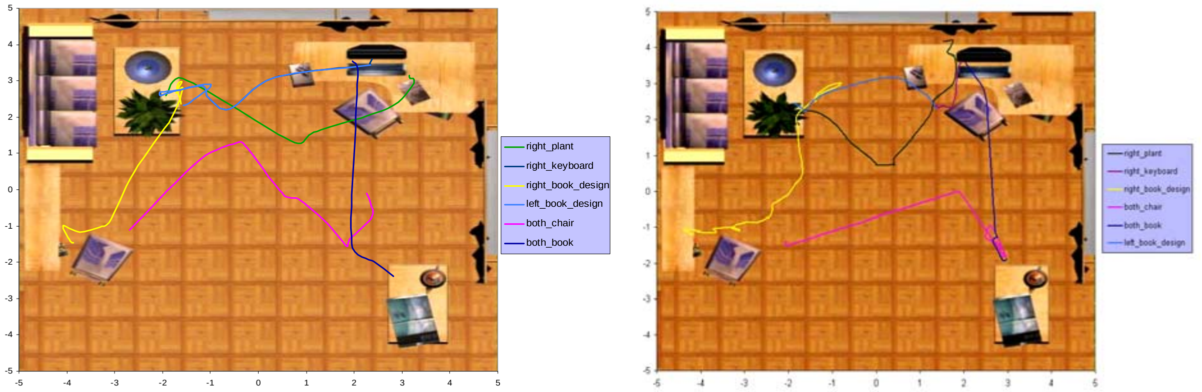
The impact on balance can be clearly seen by comparing the trajectories of the feet for subjects in both displays, Figure 3. The IPT subjects were able to smoothly walk up and down the plank turning efficiently at each end. In contrast HMD subjects stumble along the plank and seldom reach the other end without falling off. Video footage also shows stark differences. The HMD subjects continually gaze at their feet while holding their arms close to the body and often reaching for the HMD. They move reluctantly along the plank and on occasion appear close to physically falling. In contrast IPT subjects seem to gaze about one metre in front of their feet. They appeared to perceive the height and appeared to act accordingly by stretching out their arms to balance. It was interesting to observe that users of both technologies often exhibited surprise when they “fell off” the plank.

4.3 Walking along a brittle plank 4m over the floor

The results were consistent with above, however, some HMD subjects were noticeable more reluctant in their movement. One subject actually refused to do this experiment in using the HMD after experiencing a minor fall in the physical mockup.

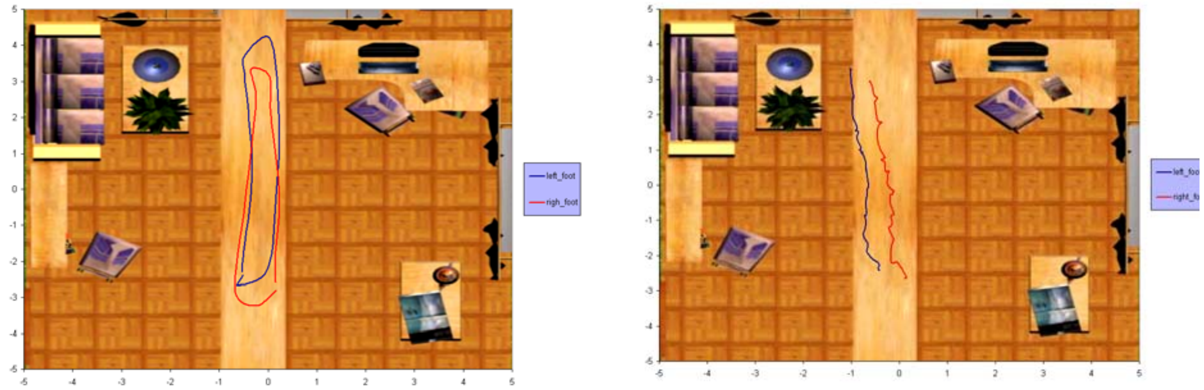
4.4 Repositioning of objects while balancing on a 0.5m high curved plank

The last task we asked participants to perform was to reposition objects while walking along a raised plank curved plank, Figure 2.



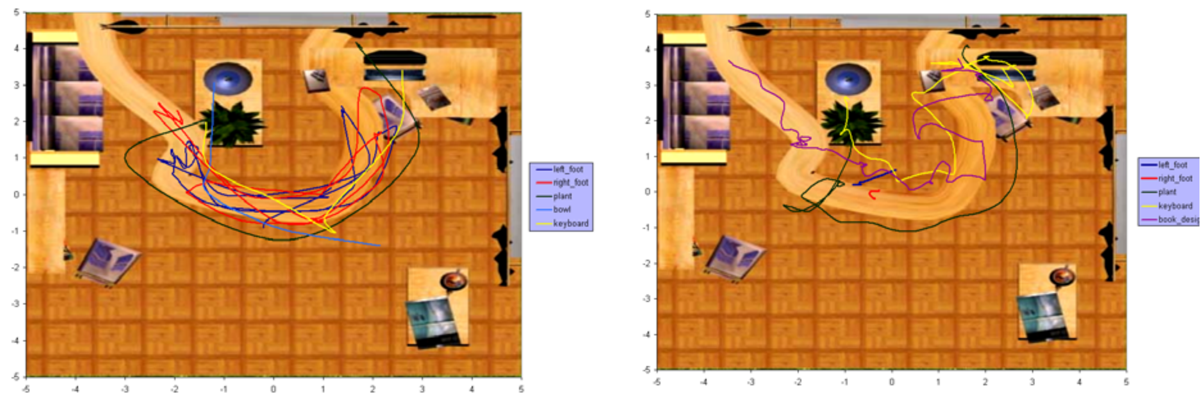
(a) (b)

Figure 3. *Path traces while repositioning of objects (a) in the IPT versus (b) HMD.*



(a) (b)

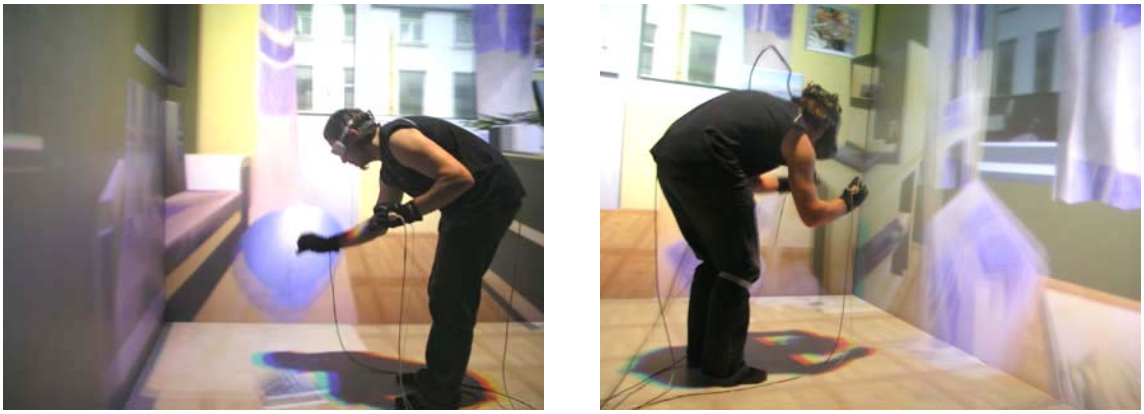
Figure 4. *Path traces while balancing on a wood plank over the room (a) in the IPT versus (b) HMD. In (a) the user did a full circle while in (b) the user fell of at the end of the plank.*



(a) (b)

Figure 5: *Path traces while repositioning of objects during balancing on a shaped plank (a) in the IPT versus (b) HMD while in (a) the user had no real problems to balance along the plank, this was much harder using a HMD.*

Both the postural and trajectory behaviours were consistent with the first two tests. IPT subjects moved efficiently and naturally, whereas HMD subjects exaggerated back and leg bends while reaching and could not balance effectively on the plank. Figure 5 contrasts typical trajectories to show that IPT subject can smoothly move around the curved plank and complete all object movement tasks whereas HMD subjects typically achieve neither.



(a) *(b)*

Figure 6. *Behaviour (body posture) during repositioning of objects (a) in the CAVE verses (b) HMD.*

5. REHABILITATION CONSIDERATIONS

The focus of this paper is on the impact of method of immersion on the naturalness of human movement. In this section we outline some wider concerns that relate to practice and placement within a clinical setting. Rehabilitation often requires teamwork between the patient and therapist. This teamwork may be hindered if the therapist is hidden from the patients view. One example of hindrance is that it hard for the therapist to demonstrate while the patient is immersed. A second is that patients may feel uncomfortable and even vulnerable if they are being watched without seeing the watcher and can be easily startled by unexpected intervention. A third hindrance is that the practitioner may not be able to relate the patient's movements to objects that are displayed on another device, for example a monitor. IPTs allow co-located users to see each other within the space, whereas HMDs require avatar representation of people. The former is more realistic but with commonly available technology, the perspective of the image is focussed on only one user and so a therapist might observe a patient reaching to one side of an object while the patient sees her hand touching it. HMDs overcome the perspective problem but introduce problems of fidelity in terms of accuracy and realism in the representation of participants.

A major consideration is practical deployment within a clinical setting. Much rehabilitation is done in cluttered spaces such as hospital, office or home. Placing the large screens of an IPT within such environment may often be impractical. Conversely, moving around a cluttered space wearing an HMD brings the danger of injury. Although IPTs have been built for less than $30K and entertainment HMDs can be purchased for much less, systems with over 110 degrees field of view regardless of orientation within the device still cost orders of magnitude more.

6. CONCLUSION

Immersive virtual reality offers considerable potential to rehabilitation as it places people in safe, easily configurable environments in which their movements can be accurately recorded. To be effective in rehabilitation it is important that the technology should not negatively impact on the naturalness of the patient's movement. We have demonstrated although IPT subjects do appear to move naturally when moving objects and balancing, HMDs can negatively affect this naturalness. In our tests, HMD subjects performed badly in balance and object placement tasks and their posture and trajectory were consistently unnatural. In contrast this IPT the movement of IPT subjects could not be distinguished from those using a physical mockup. We suspect the problems with the HMD might arise from a mismatch in proprioceptive and visual senses, brought about by not being able to see one's own body; low field of view; and stability of head set. It is possible that the mismatch in the senses can be reduced by tracking and faithfully representing all limb movement. The field of view issue and perhaps that of stability could be addressed with a state of the art HMD. Further work is needed to establish if these effects are fundamental to HMD technology or can be effectively eliminated. Specific to rehabilitation, further work is required to analyse placement and practice with regard to the characteristics of immersive displays.

Acknowledgments: The authors wish to thank the UK's Higher Education Funding Council for England for funding the equipment used in this study and also wish to thank Anthony Steed of UCL and Farshid Amirabdollahian of Salford University for useful comments.

7. REFERENCES

G Bishop and H Fuchs (1992), Research directions in virtual environments: report of an NSF Invitational Workshop, *SIGGRAPH Comput. Graph., 26*(3), pp. 153-177.

D Bowman, A Datey, Y Ryu, U Farooq and O Vasnaik (2002), *Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments*, In Proceedings of Human Factors and Ergonomics Society Annual Meeting, pp. 2134-2138.

E Burns, M C Whitton, S Razzaque, M R McCallus, A T Panter and F P Brooks (2005), *The Hand is Slower than the Eye: A quantitative exploration of visual dominance over proprioception*, In Proceedings of EEE Virtual Reality, Bonn, Germany.

J Freeman, S E Avons, R Meddis, D E Pearson and W IJsselsteijn (2000), Using Behavioral Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion-Stimuli, *Presence: Teleoperators and Virtual Environments, 9*(2), pp. 149-164.

D Hughes (2004), *Sinking in a Sea of Pixels – The Case for Media Fusion*, In Proceedings of Immersive Projection Technology Workshop.

E A Keshner and R V Kenyon (2000), The influence of an immersive virtual environment on the segmental organization of postural stabilizing responses, *Journal of Vestibular Research*, *10*(4-5), pp. 207-219.

J Kjeldskov (2001), *Interaction: Full and partial Immersive Virtual Reality Displays*, In Proceedings of IRIS24, University of Bergen, Bergen, pp. 587-600.

P D Kramer, D C Roberts, M Shelhamer and D S Zee (1998), A Versatile Stereoscopic Visual display System for Vestibular and Oculomotor Research, *Journal of Vestibular Research, 8*(5), pp. 363-379.

S Kuno, T Kawakita, O Kawakami, Y Miyake and S Wantanabe (1999), Postural Adjustment Response to Depth Direction Moving Patterns Produced by Virtual Reality Graphics, *Japanese Journal of Psychology, 49*(5), pp. 417-424.

D Manek (2004. *Effects of Visual Displays on 3D Interaction in Virtual Environments.* Virginia Polytechnic Institute and State University.

M Meehan, B Insko, M Whitton, J Frederick and P Brooks (2002), *Physiological measures of presence in stressful virtual environments*, In Proceedings of SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 645-652.

M R Mine, J Frederick P. Brooks and C H Sequin (1997), *Moving objects in space: exploiting proprioception in virtual-environment interaction*, In Proceedings of SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 19-26.

T Mulder and J Hochstenbach (2003), Motor control and learning: implications for neurological rehabilitation, In Greenwood (Ed.), *Handbook of Neurological Rehabilitation* (pp. 143-152), New York: Psychology press.

PR (2005), from www.presence-research.org

M Slater, A Steed and Y Chrysanthou (2001), *Computer Graphics and Virtual Environments: From Realism to Real-Time*: Addison Wesley Publishers.

A Steed and C Parker (2005), Evaluating Effectiveness of Interaction Techniques across Immersive Virtual Environment Systems, *Presence: Teleoperators and Virtual Environments, 14*(5), pp. (in press).

E Viirre and J Buskirk (2000), Utilization of Virtual Reality Technology in the Rehabilitation of Balance disorder Patients, *Vestibular Update, Micromedical Technologies*(24), pp. 2-4.

ICDVRAT 2006

Author Index

Author Index

©2006 ICDVRAT/University of Reading, UK; ISBN 07 049 98 65 3